ELUCIDATING THE ROLE OF SRSF1, A PROTOTYPICAL SPLICING FACTOR, IN LIVER PHYSIOLOGY

BY
WAQAR ARIF

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Biochemistry
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2020

Urbana, Illinois

Doctoral Committee:

Associate Professor Auinash Kalsotra, Chair
Professor Raven Huang
Associate Professor KV Prasanth
Assistant Professor Kai Zhang

ABSTRACT

My thesis focuses on understanding the functional role of SRSF1, a prototypical splicing factor, in the context of liver physiology. I began by generating and characterizing hepatocyte-specific SRSF1 knock-out mice models. These mice developed severe liver damage and exhibited pathology like the human disease Non-alcoholic steatohepatitis. Specifically, loss of SRSF1 results in immediate cell death and triggers a regenerative response in the liver. While SRSF1 was initially discovered as a splicing factor, studies since its discovery have revealed it plays multiple roles in RNA biology. This ranges from mRNA transport, translation regulation, non-sense mediated decay, and even processing of miRNA. Therefore, it is not surprising that SRSF1 is vital for the survival of the cell. While a deficiency in most tissues results in organ failure, the liver has the remarkable capability to circumvent the Cre-lox based knock-out of SRSF1. However, the chronic cycle of death and regeneration leads to the development of NASH-like pathology.

To understand primary changes occurring in the model, an acute SRSF1 knock-out model was developed. Insight into the molecular basis of the pathology was gained using transcriptome- and proteome-wide approaches such as RNA-Seq, eCLIP-Seq, and mass spectrometry. Extensive investigation of the molecular mechanism underlying the pathology revealed loss of SRSF1 activity leads to widespread DNA damage and defects in global translation. With the widespread damage, SRSF1 deficient hepatocytes eventually undergo necroptosis leading to failure of the liver. Moreover, these defects were also observed in siRNA mediated knockdown of SRSF1 in HepG2 cells, a human liver cancer cell line, demonstrating these effects are conserved.

ACKNOWLEDGEMENTS

It is my pleasure to acknowledge the roles of several individuals who were instrumental in the completion of my Ph.D. research.

Firstly, I would like to express my sincere gratitude to my advisor Dr. Auinash Kalsotra for the continuous support of my Ph.D. study and related research and patience. He was always motivated and continuously pushed me to strive for the best. He was always open to new ideas and gave me the opportunity to learn and grow. I would also like to thank Dr. Sayeepriyadarshini Anakk who was like a second advisor. I could always go to her for advice and I am thankful for her great insights.

Besides my advisor, I would like to thank the rest of my thesis committee: Dr. KV Prasanth, Dr. Kai Zhang, and Professor Raven Huang for their insightful comments and encouragement, but also for the hard questions which pushed me to widen my research from various perspectives.

I thank my fellow lab mates for the stimulating discussions and the too often late nights in the lab working together, and for all the fun we have had during the journey. I want to also thank my undergraduate advisor Dr. Max Funk. He taught me so much about science and experimentation in general. I would not be here today without his support. Last but not the least, I would like to thank my family: my parents and to my brothers and sister who have always been there for me.

Finally, I would like to thank my wife who has had to endure the last one and half years of me trying to finish this thesis. I could not have finished it without her continuous support and encouragement.

TABLE OF CONTENTS

List of abbreviations

A3SS	Alternative 3’ splice site
A5SS	Alternative 5’ splice site
CDS	Coding Sequence
DEG	Differentially expressed gene
FDR	False discovery rate
GFP	Green fluorescent protein
eCLIP	Enhanced UV crosslinking and immunoprecipitation
MXE	Mutually-exclusive exon
NASH	Non-alcoholic steatohepatitis
NPC	Non-parenchymal cell
RI	Retained intron
SE	Skipped exon
SRSF1	Serine Arginine Rich Splicing Factor 1
SRSF1 HKO	Hepatocyte-specific SRSF1 knock-out mice model
acSRSF1 HKO	Acute Hepatocyte-specific SRSF1 knock-out mice model
TBG	Thyroxine-binding globulin
UTR	Untranslated region

Chapter 1: Introduction

1.1 Overview

The global burden of liver-related diseases is large and steadily increasing. With the growing world-wide epidemic of metabolic syndrome, the prevalence of nonalcoholic fatty liver disease (NAFLD) is also rising (Asrih and Jornayvaz 2015). Recent reports estimate that about one-third of the United States population is affected with this disease. NAFLD is a metabolic disorder in which fat accumulates within the liver (steatosis) and can result in inflammation, hepatic injury, and cirrhosis without significant consumption of alcohol (Vernon, Baranova, and Younossi 2011). There have been significant efforts made to understand the genetic aspects that may predispose an individual to this disease. A recent study examining SNPs in patients with NAFLD revealed a significant association for pathways involved in mRNA splicing, an essential post-transcriptional processing step (Chen et al. 2013). Indeed, there are now numerous studies showing the importance of post-transcriptional gene regulatory mechanisms (PTGRM) in maintaining proper liver physiology and metabolic output. Understanding how these regulatory networks are coordinated with hepatic functions will provide valuable insight for the discovery and development of novel therapeutic targets and approaches.

RNA-binding proteins (RBP) is the broad class of proteins which act on mRNA and mediate post-transcriptional regulation. This study investigates a well characterized RBP, SRSF1, and its role in maintaining liver function and homeostasis. To study the function of this protein in-vivo, I generated and extensively characterized two variants of hepatocyte-specific knockout mice models: a transgenic model and a viral mediated

model. These mice models revealed that loss of SRSF1 in hepatocytes results in spontaneous accumulation of lipids within the cells and subsequent death.

To gain molecular insight into the mechanism of the pathogenesis, both RNA-Seq and mass spectrometry techniques were utilized to capture transcriptome and proteome changes occurring in these models. The results from this study show that SRSF1 is crucial for the stability of the genome with loss of its activity causing widespread DNA damage. Specifically, deficiency of SRSF1 results in extensive formation of R-loops which are three-stranded structures composed of a DNA-RNA hybrid and the associated non-template single-stranded DNA. Accumulation of R-loops leads to DNA double strand breaks as the cell fails to resolve these structures. Furthermore, this leads to subsequent impairment of global translation and eventual cell death by necroptosis. Reduction in global translation limits the metabolic functionality of the liver resulting in subsequent failure of the tissue. Furthermore, analysis of RNA-Seq data of SRSF1 knockdown in HepG2, a human hepatocellular carcinoma cell line, corroborates the phenotypes observed in the mice model. In brief, this study shows that the activity of SRSF1 is required for the proper functioning of hepatocytes and is vital for cellular physiology in general.

1.2 Liver Anatomy and Function

It is estimated that the liver performs nearly 500 different functions in the body, thus exemplifying its importance. It is primarily a metabolic organ, with functions ranging from but not limited to detoxification, digestion, and processing nutrients. The liver can simply be viewed as a blood filtration system both functionally and structurally. Substances and nutrients absorbed by the gastrointestinal (GI) tract are first directed to

the liver via a special venous system, known as the hepatic portal system (Gebhardt 1992). Within this first pass, nutrients present in the blood are taken up and processed to allow for systemic delivery. Furthermore, it allows the blood to be detoxified of xenobiotics and excreted before it can cause harm to the body (Sanchez and Kauffman 2010). Viewing the liver at the histological scale, its organization bears resemblance to a simple filtration system. The hepatic lobule is defined as the smallest functional unit of the liver which is comprised of the portal triad, central vein and sinusoids lined with hepatocytes bridging the two vasculatures **(Figure 1.1)** (Ben-Moshe and Itzkovitz 2019). Partially deoxygenated but nutrient rich blood from the GI enters the liver through the portal vein and is conducted into the sinusoidal capillary beds that perfuse through the liver. The sinusoid is lined with hepatocytes and a fenestrated endothelium allowing contents of the blood to exchange with the hepatocytes and vice versa. The blood then leaves the tissue through the central vein which eventually leads into the inferior vena cava.

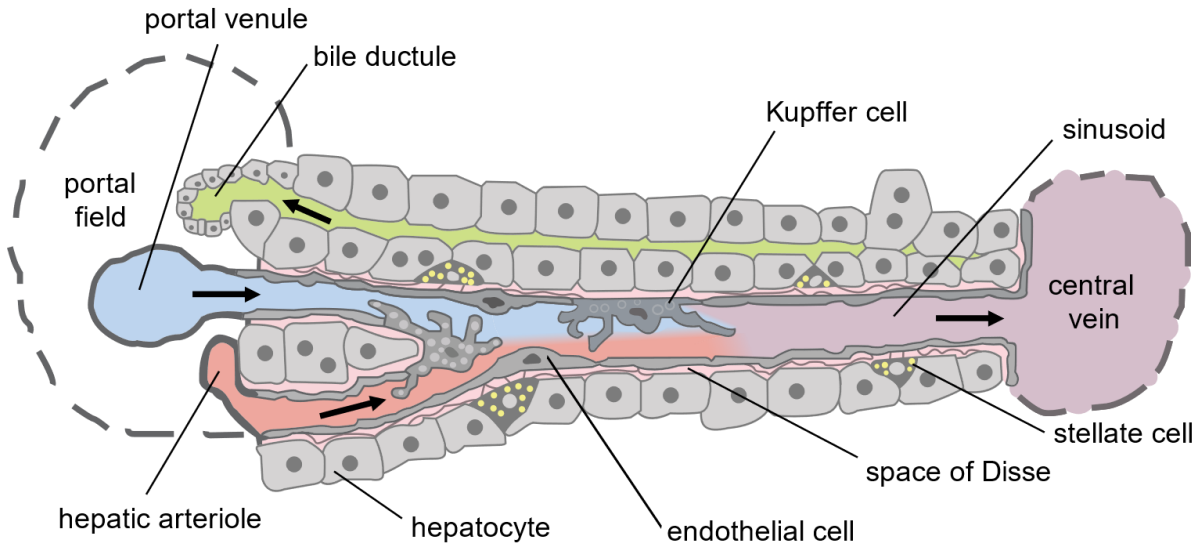


Figure 1.1: ***Organization of the Hepatic Lobule*** *(Frevert et al. 2005)*

The cell types of the liver can be classified into two broad types: parenchymal cells or hepatocytes and non-parenchymal cells. Hepatocytes constitute about 70% of the liver mass and perform the major functions of the liver such as metabolizing nutrients and drugs or secreting factors into the blood like albumin and clotting factors to name a few. Hepatocytes also produce bile which is eventually secreted into the small intestine for the digestion and absorption of lipids. While blood within the tissue flows from the portal triad to the central vein, bile flows within the bile canaliculi in the opposite direction towards the bile duct which is part of the portal triad. The non-parenchymal cells (NPC) which make up the remainder of the liver consist of cholangiocytes, sinusoidal cells, stellate cells and Kupffer cells (Si-Tayeb, Lemaigre, and Duncan 2010). Although they represent a smaller fraction, these cells are crucial for the proper development and functioning of the liver. Moreover, NPCs, such as Kupffer and stellate cells, play an important role in mounting defense response to mitigate damage from hepatotoxic stresses. Kupffer cells are the resident macrophages involved in the innate immune response and clearing away any pathogens (Liu et al. 2010). Stellate cells are normally quiescent, however, upon liver injury they become activated and release collagen to repair tissue damage (Poisson et al. 2017; U. E. Lee and Friedman 2011).

The liver plays an essential role in processing and systemic delivery of macronutrients obtained through our diet. There are two main metabolic states that a body experiences; fed (absorptive) or fasted (post-absorptive). Depending on the metabolic state, the liver behaves differently in response to an intricate feedback system involving hormones. Under fasting conditions, when blood glucose levels are low and glucagon levels are high, the liver breaks down glycogen stores into glucose and releases

it into the bloodstream. On the contrary, when glucose and insulin levels are elevated following a meal, gluconeogenesis is inhibited, and synthesis of fat, protein and glycogen are promoted in the liver (Galgani and Ravussin 2008; Patterson and Sears 2017). Metabolism in the liver is a vast and complex topic, however, for this current work, a brief overview of lipid processing and delivery is provided.

1.2.1 Packaging of Dietary Lipids and Export in the Liver

The liver plays a central role in metabolizing of lipids and exporting it throughout the body. There are three primary sources by which hepatocytes obtain triglycerides (TG): (1) *de novo* lipogenesis, (2) dietary TG entering as particles known as chylomicron remnants after feeding or (3) free fatty acids released from adipose tissue stores under fasting conditions **(Figure 1.2)**. With increased mobilization of TG into the liver, there is a simultaneous increase in its packaging and efflux into circulation. Before TG can be exported into circulation, they are first packaged into structures known as lipoprotein particles within the hepatocytes. These particles are composed of a hydrophobic core containing TG and cholesterol esters (CE) surrounded by a hydrophilic surface consisting of phospholipids (PL), free cholesterol and apolipoproteins (Zhou et al. 2015). Formation of these particles begin within the lumen of the endoplasmic reticulum (ER). The first step in the process is the translation of *apoB* transcript at the rough ER and its translocation into the lumen (Davidson and Shelness 2000). Proper folding and lipidation of this large protein into a primordial particle are facilitated by an essential factor called microsomal triglyceride transfer protein, MTTP. This factor is also responsible for forming lipid droplets within the lumen using TG present within the lumen and smooth ER. Finally, facilitated by MTTP, the lipid droplet is assembled with the primordial apoB particle resulting in a very-

low-density lipoprotein (VLDL) particle which is secreted into circulation through the Golgi **(Figure 1.2)**. Proper folding and lipid loading are necessary for the stability of apoB, otherwise the molecule becomes rapidly ubiquitinated and degraded. This can occur during conditions where lipid or MTTP levels are low (Hussain et al. 2012; Raabe et al. 1998). Although this may seem like an inefficient mechanism to regulate VLDL formation, it provides the hepatocyte the ability to rapidly adapt to changing environments. In conclusion, each step of the lipid packaging process is vital to maintaining proper functioning of the tissue and impairment at any step can result in the build of TG within the hepatocyte and development of steatosis.

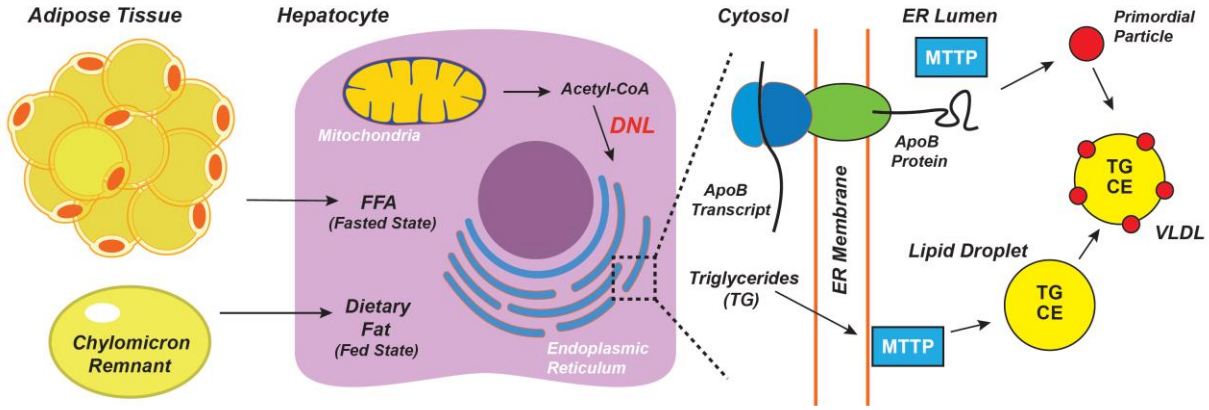


Figure 1.2: ***Assembly of Lipoprotein Particles in Hepatocytes***

Schematic providing a simplified representation of Lipoprotein particle formation within hepatocytes. Lipoprotein particles are formed within the ER lumen through the assembly of lipid droplets with newly synthesized lipoproteins. Formation of the particles is facilitated by MTTP. DNL, De novo lipogenesis; MTTP, microsomal triglyceride transfer protein; VLDL, very-low-density lipoprotein particle.

1.3 Post-Transcriptional Gene Regulation[1]

1.3.1 Introduction

Unlike prokaryotes, eukaryotic transcription and translation take place in separate cellular compartments, the nucleus, and cytoplasm, respectively. This physical uncoupling forces newly transcribed mRNA to journey into the cytoplasm before being translated into protein. It is now known that along this journey, an mRNA transcript experiences numerous interactions with various factors and undergoes extensive remodeling (Hocine, Singer, and Grünwald 2010). This diverse set of processing events includes capping, splicing, and polyadenylation and as with most biological processes, they are highly regulated, dynamic, and dependent upon the state and type of the cell. In general, these mechanisms regulate gene expression by altering the following properties of RNA: its stability, localization, translation efficiency, and nucleotide sequence. Collectively, these are referred to as post-transcriptional gene regulatory mechanisms.

1.3.2 Mechanisms of Regulation in Metabolism

Although metabolic activity is largely dictated by transcription factors and the synthesis of new mRNAs, the production of new transcripts alone does not result in greater protein expression. Each nascent transcript must first undergo a series of processing steps including splicing, chemical base modifications, and 3' end formation to be successfully translated within the cytoplasm. Thus, the process of new transcription

---

[1] **Reproduced in part with permission from**

Arif, W., Datar, G., and Kalsotra, A. Intersections of post-transcriptional gene regulatory mechanisms with intermediary metabolism. Biochim. Biophys. Acta. (2017) 1860(3), 349-362. Copyright 2017 Elsevier. The published version may be found at

https://doi.org/10.1016/j.bbagrm.2017.01.004

and protein synthesis may require hours to days to exert cellular effects. Given the delay between transcriptional activation and response, a cell can adapt to rapidly shifting metabolic conditions by mediating changes to the existing mRNA pool at the level of mRNA sequence, stability, and translation efficiency. To that end, numerous post-transcriptional regulatory mechanisms have evolved to regulate RNA processing and control gene expression. While various mechanisms of PTGRM exist, for this work only alternative splicing and translation regulation will be reviewed.

Alternative Splicing

The sequence of each mRNA largely defines its function, stability, and translation efficiency. During mRNA maturation, this sequence may be modified from the original coding gene by RNA splicing, a process by which non-coding introns are removed, and coding exons are ligated together. Importantly, the pattern in which introns and exons are spliced can vary within the same transcript in a process known as alternative splicing (AS). AS is perhaps the most well-understood regulatory step within RNA processing, and plays an important role in the expanding diversity of the eukaryotic proteome; 95% of human multi-exon genes experience a form of AS that can lead to exon skipping, alternative 3’ and 5’ splice sites, intron retention, and mutually exclusive exon inclusion (Pan et al. 2008; Y. Lee and Rio 2015). In this way, a single gene may produce multiple transcripts and proteins with varying function, stability, and localization. A specialized class of RNA binding proteins (RBP) known as splicing factors mediates the splicing of alternative exons by binding to RNA elements known as exonic splicing enhancers (ESE) and suppressors (ESS). AS is central to normal physiological function, and aberrant splicing has been implicated in several disease pathologies.

Likewise, several studies have shown that AS is integral to maintaining metabolic state and responding to physiological changes. Specifically, changes in metabolic state may alter the splicing pattern or efficiency of important metabolic regulators. Peroxisome proliferator-activated receptor gamma coactivator 1-alpha (PGC-1α) is a major transcriptional activator that has been shown to be extremely important for skeletal muscle remodeling, mitochondrial biogenesis, and angiogenesis. Variations in skeletal muscle training were found to induce a splicing switch in PGC-1α mRNA, producing the α1 and α4 isoforms. Endurance training generates the α1 isoform and leads to increases in oxidative phosphorylation genes and myosin switching. Unlike the α1 isoform, the α4 isoform of PGC-1α interacts with separate targets like Insulin-like growth factor 1 (IGF-1) and genes in the myostatin pathways to induce skeletal muscle hypertrophy in response to resistance training (Ruas et al. 2012). Furthermore, studies have outlined the role of alternative splicing regulation in cholesterol metabolism. 3-Hydroxy-3-Methylglutaryl-CoA Reductase (HMGCR), the rate-limiting enzyme in the cholesterol biosynthetic pathway, is a target of heterogeneous nuclear ribonucleoproteins A1 (hnRNP A1), a prominent regulatory protein and splicing factor. An HMGCR SNP in the binding area of hnRNP A1 was shown to promote skipping of exon 13, leading to inactivation of HMGCR activity and increased LDL-C uptake (Yu et al. 2014). Additionally, sterol loading has been demonstrated to induce the splicing and inactivation of HMGCR and LDLR by another prominent splicing factor known as Polypyrimidine Tract Binding Protein 1 (PTBP1) (Medina et al. 2011).

1.4 The SR Protein Family

The SR proteins are a family of phylogenetically conserved and structurally related RNA binding proteins which have essential roles in constitutive and alternative splicing of pre-mRNA. This family consists of 12 canonical members, all of which have a modular structure consisting of up to two N-terminal RNA-recognition motif (RRM) followed by a C-terminal RS domain which is rich in alternating serine and arginine residues (Howard and Sanford 2015; Zahler et al. 1992). These domains are functionally independent in splicing. In general, the RRM domains determine the RNA-binding properties while the RS domain mediates protein-protein interactions and is involved in orchestrating components of the spliceosome machinery. RS domains also provide nuclear localization signals for the SR protein. Depending on the strength of the localization signal of a given RS domain determines the nucleocytoplasmic shuttling activity of the associated SR protein. For instance, SRSF1 shuttles readily between the nucleus and cytoplasm while SRSF2 strictly resides within the nucleus.

Overall, SR proteins are considered to play roles in both constitutive and alternative splicing of pre-mRNA. While this statement seems counter intuitive, it can be rationalized by the modular aspect of SR protein. There is strong evidence which show that SR proteins are essential at multiple steps of spliceosome assembly (Long and Caceres 2009). However, selection of cis-acting elements by RRM domains give SR proteins to behave as alternative splicing factors. Although SR proteins play roles in housekeeping functions of the cell, an increasing number of studies are showing that nuances in their activities allow them to play regulatory roles in various processes.

1.4.1 The Role of SR Proteins in Metabolism

Recent studies are now providing evidence of SR proteins playing roles in metabolic regulation in tissue. For example, splicing factor arginine/serine-rich 10 (SFRS10) was shown to be downregulated in the liver and skeletal muscle tissues from obese patients. Lipin-1 (LPIN1), a key regulator of lipid metabolism, is a putative target of SFRS10 and exists in two major splice isoforms. Co-transfection of SFRS10 and a minigene construct containing alternative exon 6 of LPIN1 induced exon 6 skipping while SFRS10 siRNA led to an increase in exon 6 inclusion. Interestingly, the SFRS10 binding site overlaps with U1 snRNA binding sites at the 5' splice site, suggesting that binding of SFRS10 disrupts snRNA activity. Downregulation of SFRS10 favors the LPIN1B isoform, leading to an increase in lipogenesis mediated by upregulation of Srebp1c and Fatty acid synthase (FASN). Moreover, SFRS10 heterozygosity results in increased VLDL secretion and hypertriglyceridemia in mice (Dong et al. 2016).

Regulating G6PD splicing is central to the cell's ability to respond to changing metabolic states. Exon 12 of G6PD pre-mRNA contains a regulatory element that controls splicing efficiency. Under fasting conditions, upregulation of the splicing factor hnRNP K led to decreased splicing and expression of G6PD mRNA. Two C-rich motifs within exon 12 of G6PD form an ESS site leading to exon skipping and diminished splicing upon hnRNP K binding. Remarkably, these motifs also function as an ESE site for Serine/Arginine-Rich Splicing Factor 3 (SRSF3). Under nutrient-rich conditions, insulin induces SRSF3 activity through the phosphorylation of SR proteins. SRSF3 can then bind G6PD pre-mRNA and promote exon inclusion to increase G6PD expression (Walsh et al. 2013; Cyphert et al. 2013). Furthermore, it has recently been shown that SRSF3 is crucial

for hepatocyte differentiation. Deletion of SRSF3 in hepatocytes resulted in aberrant splicing of key metabolic genes necessary for glucose and lipid metabolism (Sen, Jumaa, and Webster 2013). In conclusion, there is an emerging notion that SR proteins can play roles in regulating key metabolic functions of a cell

1.5 Overview of SRSF1

SRSF1 is considered the prototypical SR protein as it was one of the first SR proteins to be identified and has been the most studied SR protein. Since its initial discovery as a factor required to complement splicing activity in S100 Hela cell extract, SRSF1 has been found to function in numerous processes within a cell (Krainer, Conway, and Kozak 1990).

1.5.1 Structure and Function of SRSF1

As mentioned previously, SR proteins contain modular domains which are linked together by flexible hinge regions. In the case of SRSF1, it contains two N-terminal RRM domains and a C-terminal RS domain **(Figure 1.3)**. As the name implies, RRM domains in SRSF1 provide it the capacity to bind RNA in a sequence specific manner. SRSF1 is involved in selection of splice sites by binding to cis-acting elements on pre-mRNA. Moreover, studies have found there to be positional dependence on SRSF1's effect on splice site selection as well.

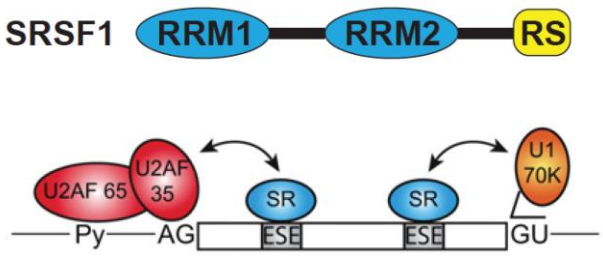


Figure 1.3: ***SRSF1 and its Role in Splice Site Selection*** (Long and Caceres 2009)

In general, binding of SRSF1 within exonic regions tend to enhance splice site selection and are therefore known as exonic splicing enhancer regions or ESE's while intronic binding sites tend to repress selection **(Figure 1.3)**. It is suggested that SRSF1 binding to exons allows it to recruit and stabilize interactions of the U1 snRNP on the 5'-splice site with U2AF on the 3'-splice site. This bridging of spliceosome factors across the exon is known as exon definition. It is thought the exon definition predominates in eukaryotic systems as genes tend to have short exons with long introns. Therefore, defining exons within a mRNA sequence is thermodynamically more favorable.

While RRM domains facilitate RNA interactions, the RS domain is involved in mediating protein-protein interactions. This domain is rich in serine and arginine residues which bestow interesting properties. The rich serine concentration of the RS domain allows for extensive phosphorylation. This modification allows for the domain to go from highly positively charged (from the arginine residues) to highly negatively charged. This change in electrostatic states allows for dynamic interactions with RNA and other proteins. Furthermore, phosphorylation of this domain is highly dynamic with studies showing SRSF1 undergoing multiple phosphorylation and dephosphorylation cycles during splicing. Interestingly, conditions where SRSF1 is kept either hyper- or hypophosphorylated, its splicing activity is inhibited. Phosphorylation also modulates the subcellular localization of SRSF1 as phosphorylation of the RS domain allows reentry into the nucleus. Taken together, the modular domains of SRFS1 are highly dynamic and confer it unique properties allowing it to have a diverse set of activities.

1.5.2 Additional Activities of SRSF1 in Various Cellular Processes

In addition to splicing, studies have shown SRSF1 to exhibit activities in essentially every aspect of gene expression regulation such as mRNA export, nonsense mediated decay, and translation to name a few (Das and Krainer 2014). A few of these activities will be discussed briefly.

Transcription and Genome Stability. There is well-documented evidence that splicing of mRNA occurs co-transcriptionally. Therefore, it is not surprising that SRSF1 has been found to interact with the RNA polymerase II complex. In fact, this interaction has been found to increase the efficiency of transcription. It is believed that co-transcriptional splicing facilitates release of RNA from the template DNA and further reduces supercoiling as the DNA can close behind the moving transcription bubble. Without co-transcriptional splicing, nascently transcribed RNA have a higher propensity to stay associated to the template DNA forming DNA-RNA hybrid structures known as R-loops (Tous and Aguilera 2007; Belotserkovskii et al. 2017). R-loops are potentially problematic as these structures are prone to DNA damage due to the exposed single strand DNA. While these structures occur naturally at some low frequency in active cells, these are not detrimental as healthy cells possess mechanisms to resolve these irregularities (Sollier and Cimprich 2015; Santos-Pereira and Aguilera 2015). However, under conditions of increased frequency of R-loop formation, a cell can easily become overwhelmed leading to DNA double strand breaks as the cell fails to resolve these structures.

There is now substantial evidence linking splicing activity to genome integrity. In fact, SRSF1 has been shown to be vital for maintaining genome stability. A 2005 study

by Li and Manley very elegantly showed that inactivation of SRSF1 in DT40 cells, a chicken B cell line, led to genomic instability. Specifically, loss of SRSF1 resulted in the accumulation of R loops which led to subsequent DNA damage by double strand breaks (X. Li and Manley 2005). This finding suggests that SRSF1 confers protection to the genome by the deleterious effects of transcription itself. SRSF1 prevents the formation of R loop by binding to nascently transcribed mRNA and preventing it from associating to template DNA.

mRNA Export. As mRNAs are co-transcriptionally spliced, a multiprotein complex known as the exon junction complex, or EJC, is deposited on the mRNA in a sequence independent manner. SRSF1 is one of the proteins that makes up the EJC (Mabin et al. 2018; Hir, Saulière, and Wang 2016). While this complex serves multiple purposes such as compacting the mRNA and providing stability, it also facilitates mRNA export to the cytoplasm. SRSF1 is a key player in mediating mRNA export. Specifically, it can act as an adaptor protein by binding directly to export factors such as TAP or Nuclear export factor 1 (NXF1) which then transports the RNP complex out to the cytoplasm (Müller-McNicoll et al. 2016; Huang, Yario, and Steitz 2004).

Translation. Of the SR proteins, SRSF1 is one of the three that displays nucleocytoplasmic shuttling activity. It was initially thought that SRSF1 may play a role in regulating translation in the cytoplasm. Soon after, a study in Hela cells showed that SRSF1 was found associated to polyribosomes in cytoplasmic extracts and that it enhanced the translation efficiency of its target (Sanford et al. 2004). Furthermore, it was shown that SRSF1 stimulates translation by promoting translation initiation of mRNA targets by suppressing 4E-BP activity, a known translation inhibitor, in a mTOR

dependent manner (Michlewski, Sanford, and Cáceres 2008). Additionally, it was found that splicing targets of SRSF1 had a greater propensity to be exhibit SRSF1 mediated translation regulation. This is one of the first examples of a factor with activities involved in coupling splicing and translation.

In conclusion, SRSF1 is a highly versatile factor with activities in numerous aspects of gene expression besides co- and post-transcriptional splicing. In addition to the functions described here, SRSF1 can regulate microRNA processing, nonsense mediated decay and nucleolar stress response. An increasing number of studies have focused on the cellular effects of SRSF1 overexpression. In fact, SRSF1 is considered a proto-oncogene as it can potentate cancer development. (Anczuków et al. 2012; Das et al. 2012). Fewer studies have investigated the effects of SRSF1 depletion. This is not surprising as SRSF1 possesses many functions its loss will result in lethality of the cell. In this study we decided to explore the effects of SRSF1 knockout in liver. As expected, loss of SRSF1 was detrimental to hepatocytes leading to robust cell death.

Chapter 2: Phenotypic Characterization of Hepatocyte-Specific SRSF1 Knock-out Mice Model

2.1 Introduction

As mentioned in section 1.4, homozygous deletion of the SRSF1 allele results in embryonic lethality in all animal models tested thus far. To circumvent this issue and study the role of SRSF1 in liver, we generated hepatocyte-specific SRSF1 knock-out mice (SRSF1 HKO) by intercrossing SRSF1 floxed mice with Alb-Cre transgenic mice. Since there have been no previous study of regarding SRSF1's role in liver, characterization of the phenotype was performed using histological and serum analysis at various time-points in development. Previous studies using the Albumin-Cre transgenic mice report Cre to be homogenously active around embryonic day 18.5. For this reason, tissues were harvested at early time points up to 10 days after birth as well as later time points in development starting at 5 weeks. Histological analysis provides valuable insights into microscopic changes occurring of the tissue anatomy. Various stains have also been performed to allow for understanding different aspects of the tissue architecture. Similarly, analysis of specific serum parameters can provide useful information regarding the status of the liver.

2.2 Material and Methods

The following sections provides details of the materials and methodology utilized for the studies performed in this chapter. For details on additional methodology including western blot analysis, histological staining, immunohistochemistry, and immunofluorescence staining, refer to *Section 9.1: Supplementary Protocols.*

2.2.1 Generating Hepatocyte-specific SRSF1 Knockout Mice Model

Mice with floxed alleles for SRSF1, developed by Dr. Xiang-Dong Fu and available from The Jackson Laboratory, were bred with mice expressing the Cre-recombinase driven by the Albumin promoter to obtain hepatocyte-specific SRSF1 knockout (SRSF1 HKO) mice **(Figure 2.1)** (X. Xu et al. 2005; Postic et al. 1999). The mice were housed on a standard 12-hour-light/dark cycle and fed a normal chow diet (2918 Envigo Teklad). The mice were fasted for 9 to 10 hours prior to harvesting tissues for subsequent analysis. National Institutes of Health (NIH) guidelines for use and care of laboratory animals was followed and all the experiments were approved by the Institutional Animal Care and Use Committee at the University of Illinois at Urbana-Champaign.

2.2.2 Blood Collection and Serum Chemistry Analysis

Mice were fasted about 10 hours prior to blood collection. Blood was collected from mice by bleeding them from the retro-orbital venous sinus using EDTA coated blood collecting capillaries. Mice were temporarily anesthetized using isoflurane before collecting blood. For non-terminal procedures about 200 µL was collected otherwise about 600 µL was collected. Blood from the capillary was transferred into BD microtainer tubes and then centrifuged according to manufacture’s protocol to separate the serum. Serum was then transferred into 1.5 mL microcentrifuge tubes and snap frozen in liquid

nitrogen before storing at -80 °C. Serum cholesterol, triglycerides, ALT and AST activity were measured using colorimetric assay kits provided by Infinity (Thermo Scientific). Serum bilirubin, both direct and indirect, were measured using the Sigma Bilirubin Assay kit (MAK126 Sigma) according to the manufacturer's protocol.

2.2.3 Glucose Tolerance Test in Mice

Mice were fasted overnight and injected with D-glucose at 2g/kg intraperitoneally. Blood glucose concentrations were measured at 0, 15, 30, 45, 60 and 120 min after glucose injections using the One Touch glucose meter. Blood for measurement was obtained from the tip of the tail after clipping.

2.2.4 Generating Diet Induced Hepatic Steatosis Mice Model

Starting at 8 weeks of age, C57B/6J mice were fed 60% high fat diet (Envigo Teklad) for 8 weeks to mimic obese conditions. Mice were sacrificed at the end of the experimental regimen. Liver and white adipose tissues were collected for gene expression, western blotting, and histological analysis.

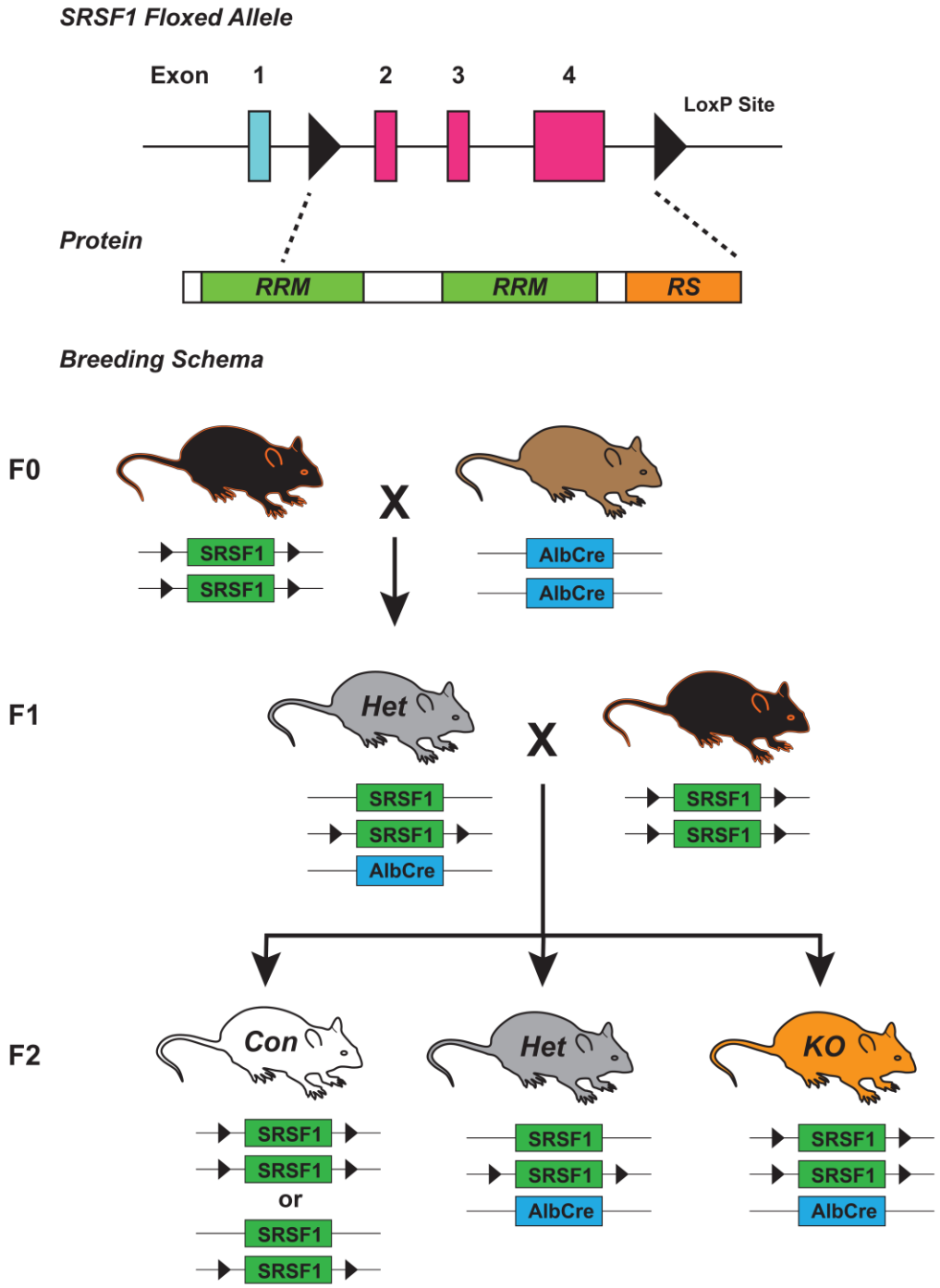


Figure 2.1: ***Generating Hepatocyte-Specific SRSF1 KO Mice Model***

The SRSF1 Flox mice have LoxP sites flanking exons 2 through 4 resulting in significant removal of the open reading frame. The breeding schema was used to generate SRSF1 HKO mice along with Controls (Con). Het animals were not used for this study.

2.3 Results

2.3.1 Loss of SRSF1 in Hepatocytes Results in Immediate Development of Steatosis

SRSF1 HKO mice exhibit immediate and robust development of lipid accumulation within the liver tissue. This is apparent from the gross appearance of the liver tissue 10 days after birth **(Figure 2.2 A)**. Compared to controls, SRSF1 HKO livers at this stage are strikingly yellow in color, signifying an accumulation of fat within the tissue. Mice at this age also weigh less than littermate control mice **(Figure 2.2 B)**.

Histological staining was performed on tissue sections at early timepoints in development to gain further insight into the pathology **(Figure 2.3)**. Hematoxylin and Eosin staining was performed to observe general morphological changes occurring in the tissue and Oil Red O staining was performed to qualitatively assess the extent of lipid accumulation within the tissue. From the staining it is apparent that starting at 6 days after birth, hepatocyte death can be observed in SRSF1 HKO livers. Hepatic cell death is observed at later timepoints as well. With regards to hepatic lipid levels, both control and SRSF1 HKO show significant accumulation of lipids in the tissue up to 6 days of age. However, after 6 days, control livers show severe reduction in lipids within the tissue whereas SRSF1 HKO have persistent accumulation **(Figure 2.3)**.

To assess the knockout efficiency of SRSF1 in this mice model, western blot analysis was performed on isolated hepatocytes. Interestingly, we find about a 60% reduction in SRSF1 levels in hepatocytes 10 days after birth **(Figure 2.2 C)**. However, previous reports have shown that the *albCre* transgenic driver results in nearly complete conversion of hepatocytes by postnatal day 3 using a Cre-dependent fluorescent reporter (Postic et al. 1999; Weisend et al. 2009). Reduction of SRSF1 observed in our model can

be a rationalized with the following possibilities; (1) there is a 60% reduction in SRSF1 levels in all hepatocytes or (2) about 60% of hepatocytes are have SRSF1 knocked out while the remaining 40% are still expressing SRSF1 at wildtype levels.

To distinguish between the two possible mechanisms, an immunofluorescent staining was performed on tissue sections from control and SRSF1 HKO livers **(Figure 2.4)**. An advantage of this technique is that it allows for the probing of multiple factors. Therefore, in addition to probing for SRSF1, the tissue section is also probed for HNF4A, a hepatocyte specific marker. This allows for efficient identification of hepatocytes within a field and subsequent quantification of SRSF1 within those cells. Quantification is performed using custom ImageJ scripts **(Figure 2.5 A)**. Greatest knockout of SRSF1 within hepatocytes is observed 6 days after birth. After this point, the liver tissue begins to slowly repopulate with SRSF1 expressing hepatocytes. Tissue sections were also stained for Ki67, a cell division marker, to allow for comparison of proliferation between controls and SRSF1 HKO **(Figure 2.4)**. Quantification of proliferating hepatocytes reveal that the fraction of proliferating cells decreases at 6 days in SRSF1 HKO, but then increases afterwards **(Figure 2.5 B)**.

2.3.2 Development of NASH-like Pathology in SRSF1 HKO Mice

While livers of SRSF1 HKO mice can circumvent knockout of SRSF1 in the liver, these mice still develop severe pathology. Gross inspection of liver tissue at the five-week and three-month timepoint show steatosis (paler color), however, the severity is far less in comparison to ten-day livers. At 5 weeks, western blot analysis shows about 2-fold reduction in abundance compared to controls **(Figure 2.6 C)**. Characterization of SRSF1 HKO shows a growth retardation phenotype with elevated ALT and AST levels starting at

five-week timepoint signifying liver damage **(Figure 2.6 A-B and 2.8 A)**. However, by six months these mice catch up in growth to controls and exhibit no difference in ALT and AST levels. Earlier timepoints also have decreased levels of cholesterol and triglycerides at the 5 weeks and 3 months timepoints **(Figure 2.8 A).** Furthermore, at 8 weeks, glucose tolerance testing of control and SRSF1 HKO mice do not show any significant differences **(Figure 2.8 B)**

This is in agreement with the hypothesis that the hepatocytes over time are bypassing the knockout. Liver-to-body weight ratio in SRSF1 HKO do not show a significant difference in comparison to control. However, gonadal fat-to-body weight ratio shows significant reduction at both the 5 week and 3-month time point **(Figure 2.7)**. Histological analysis at these later timepoints reveal development of NASH-like pathology such as inflammation and fibrosis **(Figure 2.9)**. At the five-week timepoint, severe damage of the liver is seen from the H&E staining with infiltration of inflammatory cells and necrosis of hepatocytes. However, at later timepoints the inflammation and necrosis reduce. Oil Red O staining reveals lipid accumulation within the tissue at both 5 week and 3-month timepoint, however, by 6 month the lipid levels decrease. Fibrosis is also present by 3 month and develops to bridging fibrosis by 6 months **(Figure 2.9)**.

2.3.3 Cell Death in SRSF1 HKO does not occur via Apoptosis

Significant cell death is observed in SRSF1 HKO model at earlier timepoints. To determine if the cell death occurs via apoptosis, TUNEL staining was performed on tissue sections **(Figure 2.10 C)**. SRSF1 HKO tissue sections do not show any TUNEL staining signifying hepatocytes do not undergo apoptosis. It is most likely that loss of SRSF1 causes cell death by a necrosis or necroptosis.

2.3.4 Repopulation of SRSF1 Expressing Hepatocytes in Adult SRSF1 HKO

To evaluate the extent of repopulation of the liver with SRSF1 expressing hepatocytes at adult stages, IF analysis was performed at 5 weeks and 3 months timepoints **(Figure 2.10 A)**. Quantification reveals that livers of SRSF1 HKO mice at 5 weeks are still in the process of repopulating with nearly 60% of hepatocytes expressing SRSF1. This corroborates with the western blot analysis shown previously in *Figure 2.6 C*. By 3 months, the tissue is completely repopulated by SRSF1 expressing hepatocytes **(Figure 2.10 A)**. This repopulation is most likely a result of shutting off of Cre expression. To test this hypothesis, Cre expression was measured at both 5 weeks and 3 months using qRT-PCR analysis. Compared to controls (AlbCre +/-), SRSF1 HKO livers have a remarkable downregulation of Cre expression **(Figure 2.10 B)**.

2.3.5 SRSF1 HKO by 1.5 Years Develop Nodules and Severe Inflammation

SRSF1 HKO mice aged to 1.5 years tend to develop growths on the liver **(Figure 2.11 A)**. Further inspection of these growths using histological analysis reveals there is a striking infiltration of inflammatory cells and leukocytes within the tissue. These infiltrations also correspond to the nodular growths on the tissue **(Figure 2.11 B)**. This infiltration also extends from portal triads. This inflammation is comparable to the inflammation observed in autoimmune hepatitis. Further IF imaging was performed to determine if the observed damage is a result of SRSF1 deficiency. However, IF imaging reveals that 1.5-year-old SRSF1 HKO liver tissue still contain SRSF1 expressing hepatocytes **(Figure 2.11 C)**. One possible explanation of the extensive inflammation observed in aged SRSF1 HKO is a result of some inherent viral infection present within

the mouse facility. However, if this were the case, frequency of inflammation should be equivalent between control and SRSF1 HKO mice. Another possible explanation for the inflammation is that SRSF1 HKO mice exhibit an increased propensity to develop an autoimmune reaction to the increased aneuploid hepatocytes. While it has not been shown that SRSF1 HKO exhibit an increase in the number of aneuploid hepatocytes, this is most likely the case based on previous reports and the loss of Cre expression. Further investigation is necessary to distinguish the cause of the increased inflammation in aged SRSF1 HKO mice.

2.3.6 SRSF1 Levels are not Regulated by Hepatic Lipid Content

To check if SRSF1 levels are regulated by hepatic lipid content, SRSF1 levels were measured in a steatosis model. Adult mice were fed a high-fat diet for 8 weeks to allow for the accumulation of lipids within the liver. Brief characterization of the model reveals that the mice develop severely fatty livers **(Figure 2.12 A-C)**. This is evident from the gross images of the livers collected from this model after 8 weeks of high-fat diet treatment. Livers from the high-fat fed mice are strikingly yellow in appearance due to the significant buildup of lipids within the tissue **(Figure 2.12 C)**. Western blot analysis performed on isolated hepatocytes show there is no significant change in SRSF1 in this hepatic steatosis mice model **(Figure 2.12 D)**. With these findings, it can be concluded that abundance of SRSF1 is not responsive to hepatic lipid levels.

2.4 Conclusions

It is now well-known that post-transcriptional gene regulation contributes significantly to maintaining normal tissue physiology. To study the role of the prototypical splicing factor, SRSF1, within the liver, a hepatocyte specific SRSF1 knockout model was developed. Characterization of SRSF1 HKO mice early in development reveal robust accumulation of lipid within the liver **(Figure 2.2 A and 2.3)**. This is followed by significant cell death, inflammation, and eventual development of fibrosis **(Figure 2.9)**. The pathology seen in this model closely follows that of the human liver disease known as Non-alcoholic steatohepatitis. However, further inspection of the phenotype reveals that the liver damage observed in SRSF1 HKO is a result of continuous cell death and regeneration. This is evident from the immunofluorescent (IF) and western blot analysis performed on the tissue at various time-points in the SRSF1 HKO model.

Immunofluorescent (IF) analysis on tissue sections showed the greatest knockout of SRSF1 in SRSF1 HKO mice occurs at post-natal day six. To circumvent the loss of SRSF1, the liver begins regenerating and repopulating with SRSF1 expressing hepatocytes as soon as post-natal day eight. This immediate repopulation of the liver with SRSF1 expressing cells demonstrates the necessity of SRSF1 for hepatocyte survival. Without SRSF1, hepatocytes cannot function properly and eventually die. Repopulation of the liver in this model is most likely a contribution from two separate sources; 1) the proliferation of the small fraction of hepatocytes which escaped knock-out of SRSF1 and 2) transdifferentiation of biliary epithelial cells (BEC) into hepatocytes. However, most hepatocytes originating from transdifferentiating BECs would undergo knock-out of SRSF1 as the cell begins expressing *Albumin* along with the Cre recombinase. This cycle

of hepatocyte proliferation and cell death due to SRSF1 knock-out continues over a span of time until the tissue slowly repopulates itself with hepatocytes resistant to SRSF1 knock-out. This is evident from the IF staining at later timepoints which show that the fraction of hepatocytes expressing SRSF1 is increasing with time. While this may seem unlikely, this type of phenomena has been previously described and is possible due to a unique feature of the liver (Duncan et al. 2012; Duncan 2013).

One of the most remarkable and often overlooked attributes of the liver is that over half of the mature hepatocytes are aneuploid. While these mature hepatocytes exhibit multiple chromosomal gains and losses, these cells are functional and can undergo mitosis. It is hypothesized that this feature of the liver has evolved as an adaptive mechanism to hepatoxic stresses. The widespread genetic heterogeneity provides a fitness landscape allowing selection for the fittest hepatocytes to persist against various insults. Support for this hypothesis was eloquently demonstrated by *Duncan et al.,* utilizing a genetic liver disease model for hereditary tyrosinemia type I (Duncan et al. 2012). This model is deficient of fumarylacetoacetate hydrolase (FAH), a critical enzyme in the metabolism of tyrosine, which results in toxic accumulation of fumarylacetoacetate. However, buildup of this toxic metabolite can be prevented by loss of an upstream enzyme known as homogentisic acid dioxygenase (HGD). To test if the liver could utilize adaptive aneuploidy to resist injury, *Duncan et al.* cleverly developed mice deficient for FAH and heterozygous for a mutation in HGD. Remarkably, livers in these mice developed regenerating nodules enriched with hepatocytes exhibiting a loss of chromosome 16 which contains the locus for *Hgd*. Specifically, chromosome 16-specific aneuploidy was approximately found in 50 % of the hepatocytes making the mice resistant

to injury. Similarly, this phenomenon of adaptive aneuploidy is occurring within the SRSF1 HKO to ensure survival of the liver.

Parallels exist in the model used by *Duncan et al*. to investigate adaptive aneuploidy and the SRSF1 HKO model. While loss of the *Hgd* locus is protective in their model, loss of the *AlbCre* locus is protective in SRSF1 HKO. Loss of the Cre recombinase would prevent knock-out of SRSF1 and ensure survival of the hepatocyte. This is strongly supported by the results obtained from the SRSF1 IF staining of the liver tissue sections at various timepoints in the SRSF1 HKO **(Figure 2.4 and 2.7)**. Repopulation of SRSF1 expressing hepatocytes is seen as early as post-natal day eight and continues to slowly repopulate until it reaches completion by three months. Greatest knock-out is observed at post-natal day six, however, even at this timepoint there are a small fraction of hepatocytes which are expressing SRSF1. These are most likely aneuploid cells which are deficient of the *AlbCre* locus and begin to proliferate in the early stages to repopulate the tissue. However, to overcome the severe hepatocyte deficiency, BECs begin transdifferentiating into hepatocytes. This is evident from the SRSF1 IF staining in the SRSF1 HKO which show population of SRSF1 expressing cells in proximity to the portal triad. However, at the early timepoints most of these cells contain the *AlbCre* locus and will result in knock-out of SRSF1 and eventual cell death. This cycle will continue until the liver is primarily consists of aneuploid cells which lose expression of Cre and thus are able to express SRSF1. Indeed, measurement of Cre mRNA expression using qRT-PCR at 5 weeks and 3 months revealed significant reduction in Cre expression **(Figure 2.10 B)**. The chronic cycle of cell death and regeneration leads to the NASH-like pathology observed in this model. However, as the liver is repopulated with an increasing number

of SRSF1 knock-out resistant hepatocytes, the inflammation, damage, and lipid accumulation subside as well. Moreover, by eight weeks metabolic tests such as glucose tolerance test and fasting glucose levels do not show any significant differences between control and SRSF1 HKO mice **(Figure 2.7 and 2.8)**.

Finally, to confirm whether hepatic lipid content can regulate SRSF1 expression, levels of SRSF1 were checked in a diet induced steatosis mice model **(Figure 2.12)**. Mice were put on a high-fat diet regime for eight months which resulted in significant accumulation of fat within the liver. Western blot analysis on hepatocytes isolated from these fatty livers shows that there was no difference in SRSF1 levels in comparison to the control chow-fed animals. We can therefore conclude that levels of SRSF1 in hepatocytes are not dynamically regulated by hepatic lipid content. To gain insight into the molecular mechanisms causing the pathology observed in SRSF1 HKO mice, RNA-Seq analysis was used to investigate transcriptome-wide changes in expression and splicing.

2.5 Chapter Figures

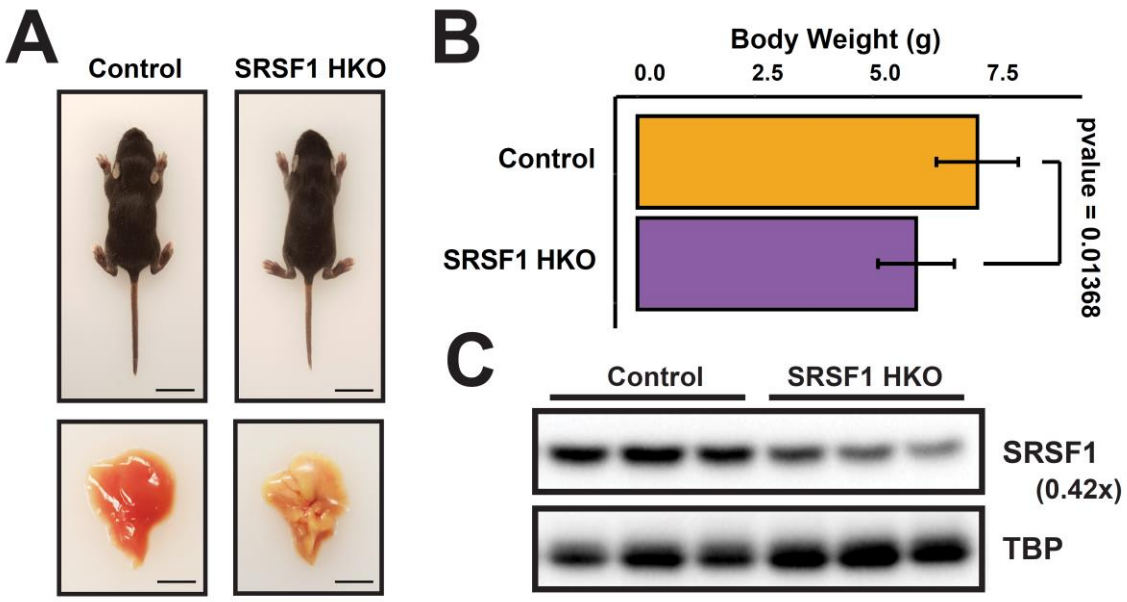


Figure 2.2: ***SRSF1 HKO Mice Models Develop Steatosis***

(A) Gross images of the Control and SRSF1 HKO mice 10 days after birth. Development of severe steatosis is apparent at this timepoint. (B) Body weight measurements shows significant reduction in body weight in SRSF1 HKO (n = 6) (C) Abundance of SRSF1 was measured in hepatocytes from SRSF1 HKO. A significant reduction in SRSF1 abundance is seen in SRSF1 HKO.

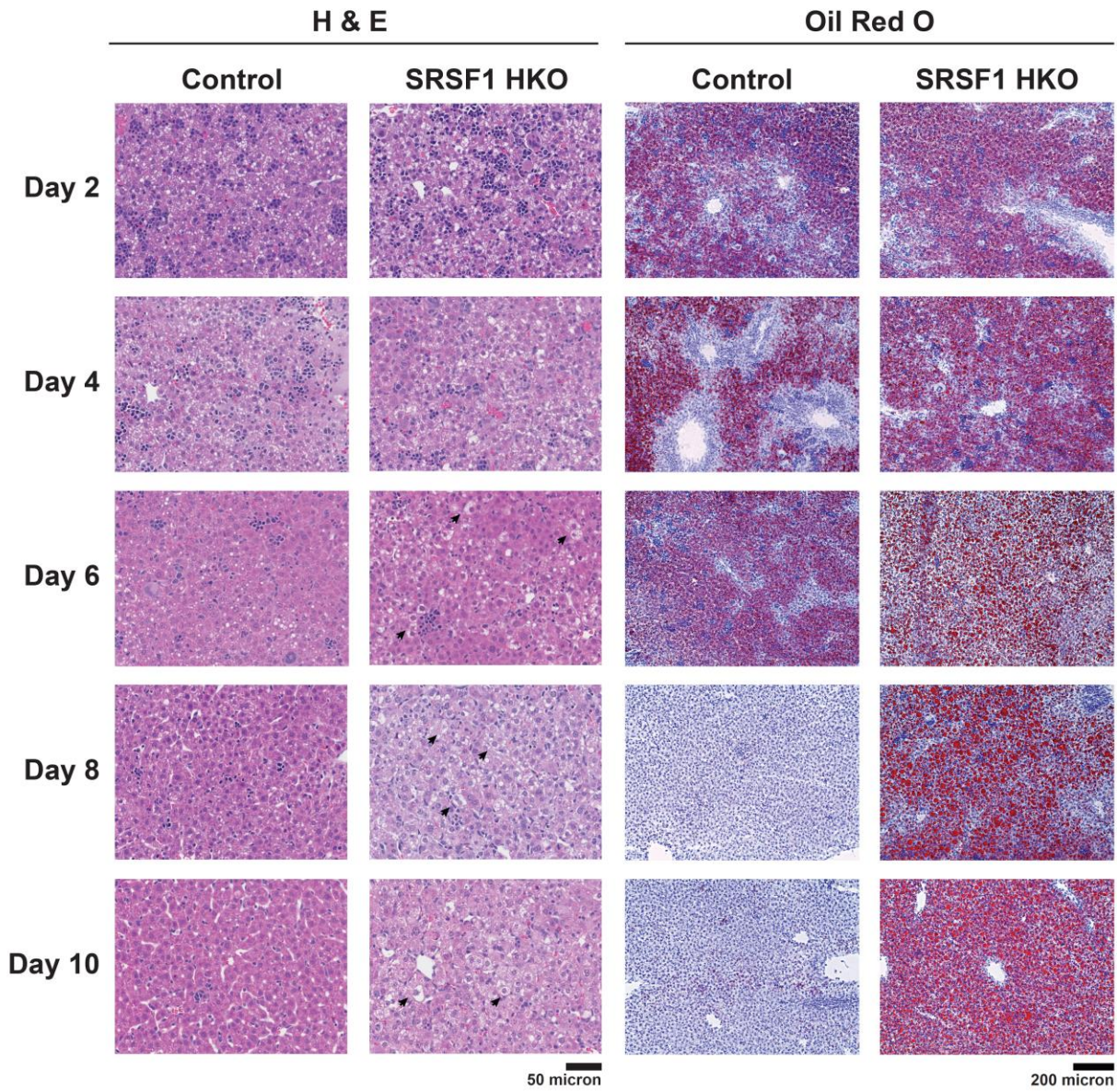


Figure 2.3: ***SRSF1 HKO Livers Develop Severe Lipid Accumulation Early in Development***

To qualitatively assess the morphological changes within the tissue of SRSF1 HKO mice, both H&E and Oil Red O staining was performed. H&E staining allows for assessment of general tissue architecture while Oil Red O allows for the assessment of lipid accumulation.

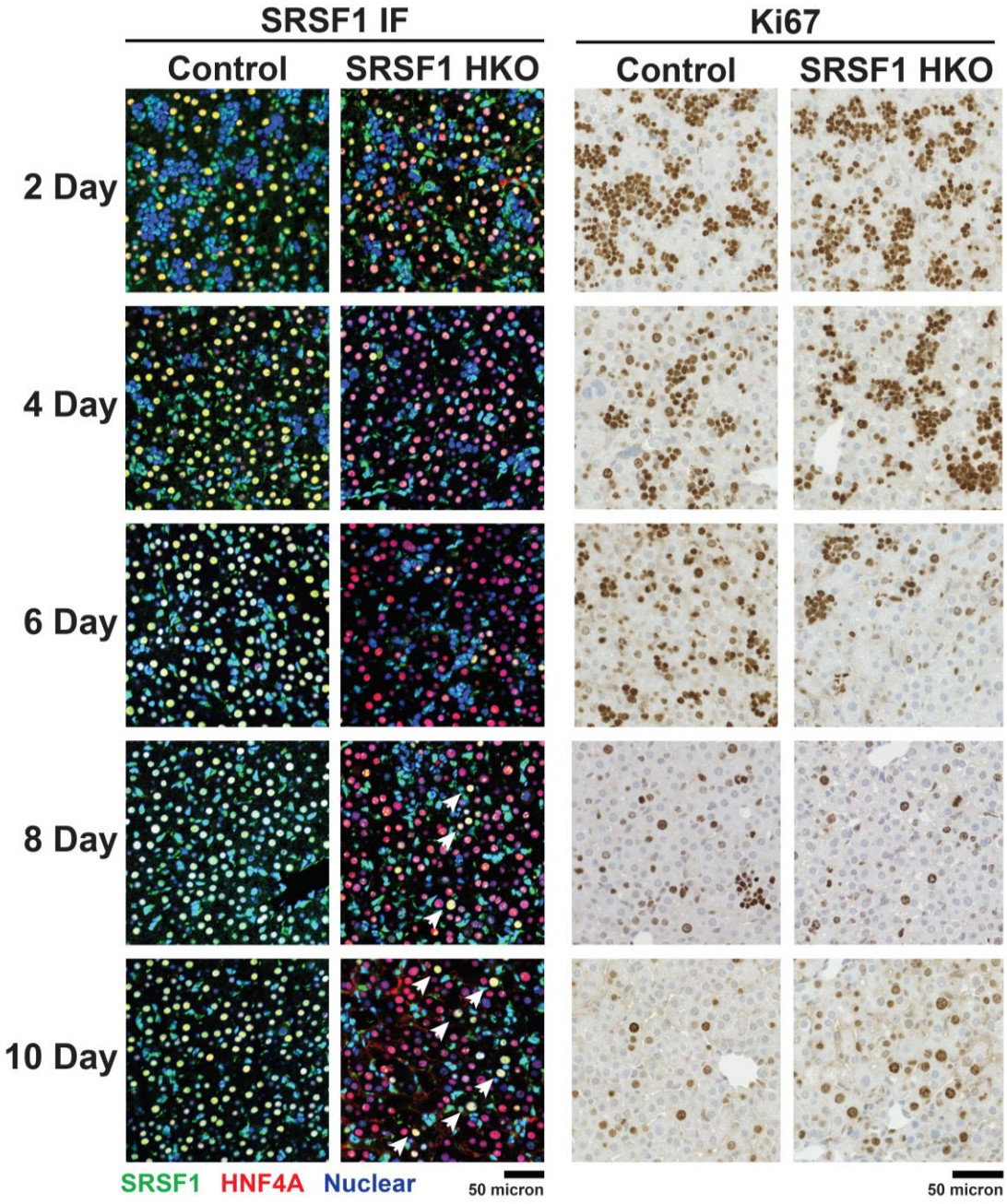


Figure 2.4: ***SRSF1 HKO Livers Begin Repopulating with SRSF1 Expressing Hepatocytes Immediately After Knock-out***

To assess the extent of knockout both spatially and temporally in the SRSF1 HKO livers, immunofluorescence staining was used on tissue sections (n = 3, 6 fields) at various timepoints (SRSF1, green; HNF4a, red; Nuclei, blue). Immunohistochemical staining for ki67 was also performed at these timepoints to assess the extent of hepatocyte proliferation.

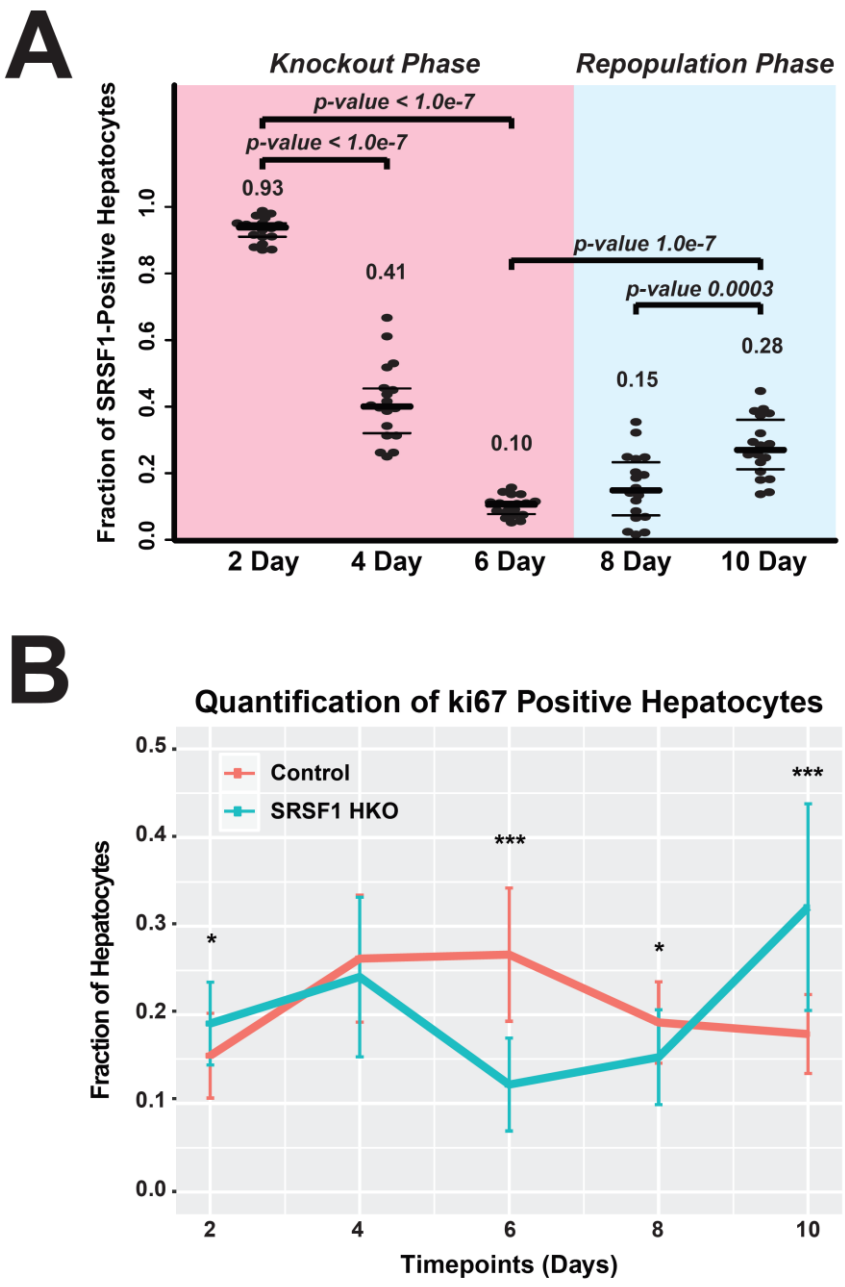


Figure 2.5: ***Quantification of SRSF1 Expressing Hepatocytes and Proliferation at Early Timepoints***

Quantification of hepatocytes expressing SRSF1 (signal intensity > 10% of control levels) and (B) fraction of hepatocytes proliferating using KI67 as the marker.

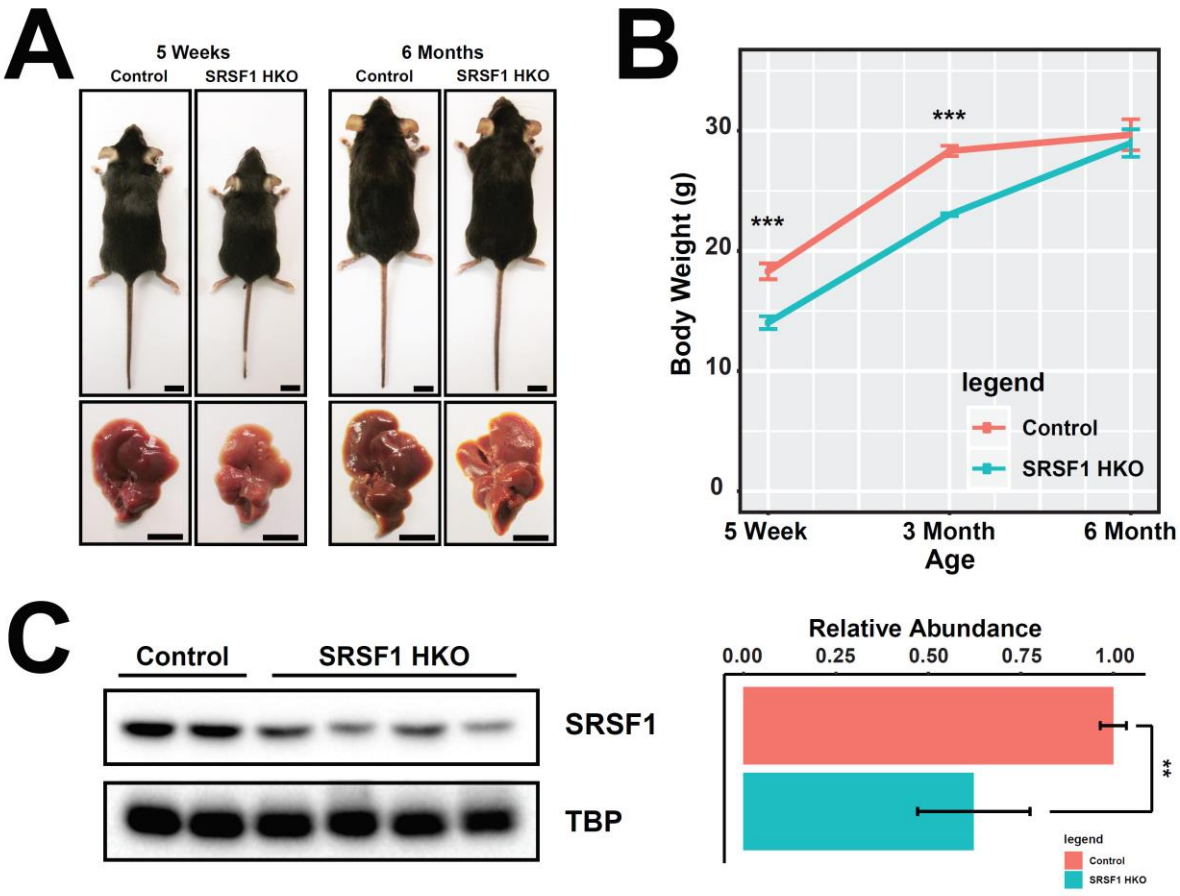


Figure 2.6: ***Characterization of SRSF1 HKO Mice during Adult Development***

(A) Gross appearance of control and SRSF1 HKO mice and livers at 5 weeks and 6 months timepoint. Scale bar = 1 cm. (B) Body weight growth curve of controls and SRSF1 HKO mice. (C) Relative SRFS1 protein abundance quantification in SRSF1 HKO mice at 5 weeks using western blot analysis.

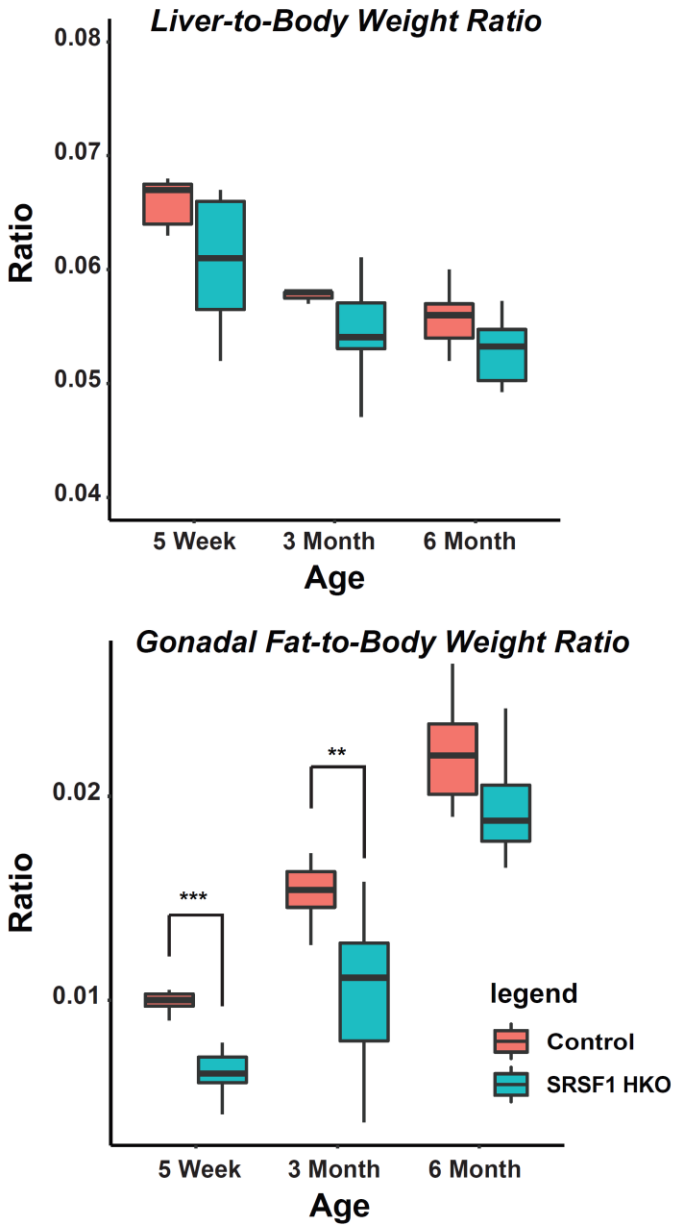


Figure 2.7: ***Liver- and Adipose-somatic index values for SRSF1 HKO mice***

Liver-to-body weight and Gonadal fat-to-body weight ratios in control and SRSF1 HKO mice during adult development. Liver-to-body weight ratios are not significantly different between control and SRSF1 HKO. However, gonadal fat-body weight ratio shows significant decrease at earlier timepoints.

A

Serum Parameter	Age	Control	SRSF1 HKO	P-Value	Direction
Blood Glucose (mg/dL)	5 Week	118.7 ± 3.8	121.7 ± 6.8	0.2925	N.S.
	3 Month	129.4 ± 4.6	118.7 ± 9.3	0.0242	Down
	6 Month	138.7 ± 19.7	125.0 ± 20.6	0.2186	N.S.
ALT (U/L)	5 Week	19.2 ± 8.5	231.1 ± 58.7	0.0001	Up
	3 Month	13.9 ± 5.6	27.8 ± 12.1	0.0474	Up
	6 Month	9.14 ± 5.2	15.4 ± 5.6	0.1005	N.S.
AST (U/L)	5 Week	43.1 ± 31.2	225.2 ± 39.4	0.0001	Up
	3 Month	25.7 ± 7.8	89.8 ± 17.0	0.0002	Up
	6 Month	41.1 ± 10.8.5	42.5 ± 18.5	0.9018	N.S.
Triglyceride (mg/dL)	5 Week	80.4 ± 13.7	65.5 ± 2.9	0.0277	Down
	3 Month	143.7 ± 10.6	100.5 ± 12.6	0.0000	Down
	6 Month	115.9 ± 10.7	128.9 ± 22.1	0.2020	N.S.
Cholesterol (mg/dL)	5 Week	119.2 ± 22.7	90.2 ± 9.9	0.0343	Down
	3 Month	96.7 ± 22.4	120.6 ± 17.3	0.0350	Down
	6 Month	103.1 ± 8.0	121.5 ± 21.0	0.0612	N.S.

B

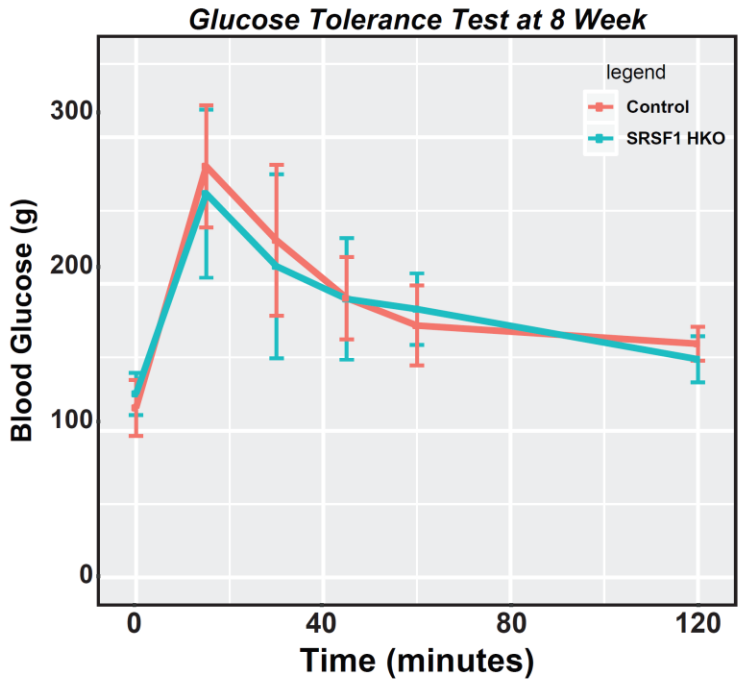


Figure 2.8: ***Serum Metabolic Profile and Glucose Tolerance Test in SRSF1 HKO***

(A) Serum parameters in control and SRSF1 HKO mice during adult timepoints. ALT and AST show significant increase at earlier timepoints while cholesterol and triglycerides are significantly decreased at earlier timepoints. All parameters normalize by 6 months of age. Values are presented as mean ± standard deviation. (B) Glucose tolerance test of control and SRSF1 HKO mice at 8 weeks. No significant difference observed.

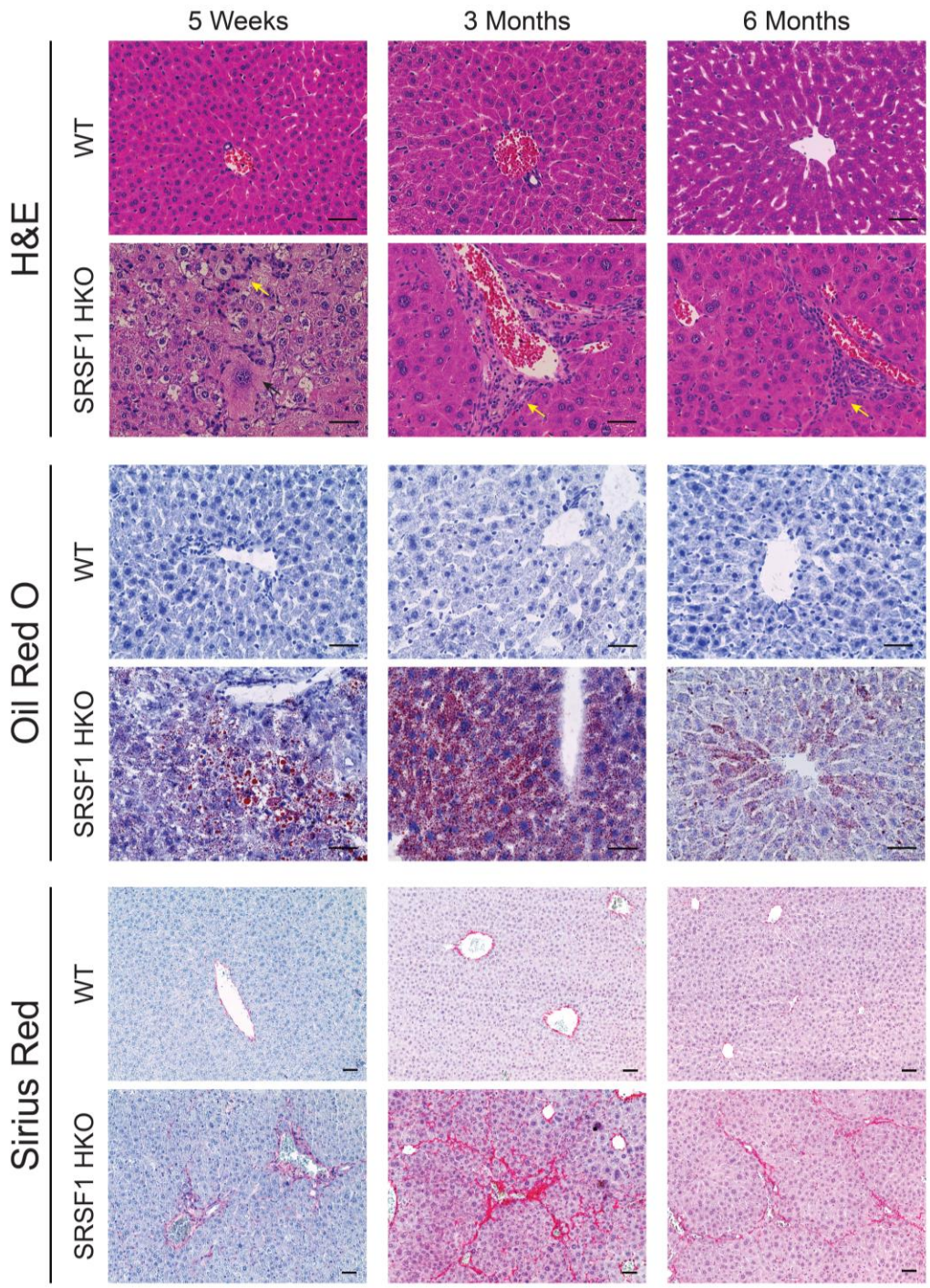


Figure 2.9: ***Histology of SRSF1 HKO at Various Timepoints***

Histology of SRSF1 HKO at 5 weeks, 3 months, and 6 months. H&E shows inflammation at 5 weeks. SRSF1 HKO have increased lipid accumulation at earlier timepoints. SRSF1 HKO develop fibrosis as seen by Sirius Red by 3 months. Scale bar = 100 micron.

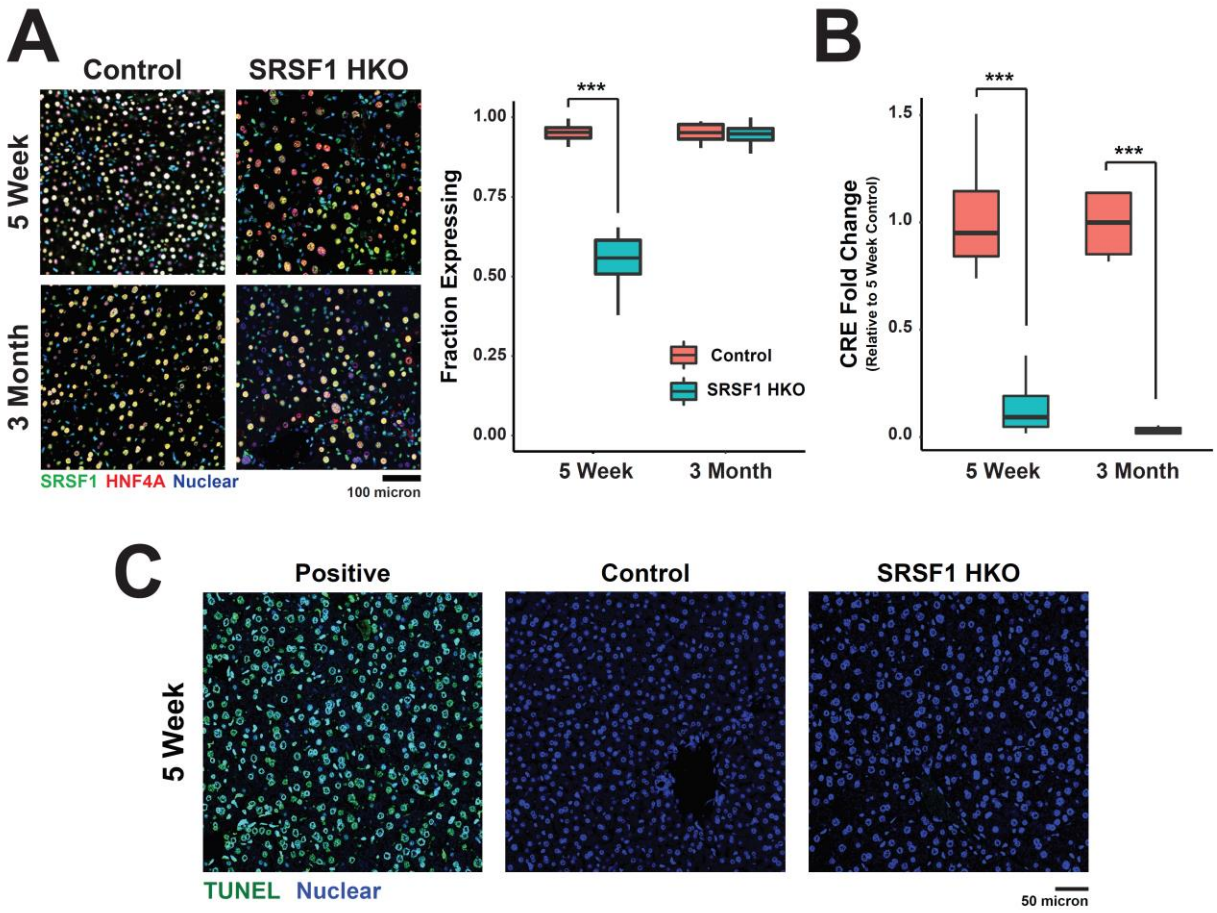


Figure 2.10: ***Regenerating SRSF1 HKO livers evade knockout by shutdown of CRE***

(A) SRSF1 expressing hepatocytes at 5 weeks and 3 months timepoint shows progressive repopulation. At 5 weeks nearly 60% of the hepatocytes are expressing SRSF1 with nearly complete repopulation by 3 months. (B) Expression of Cre by qRT-PCR shows significant reduction of levels in SRSF1 HKO. (C) SRSF1 HKO do not have any TUNEL staining signifying the cell death occurring in this model is not via apoptosis.

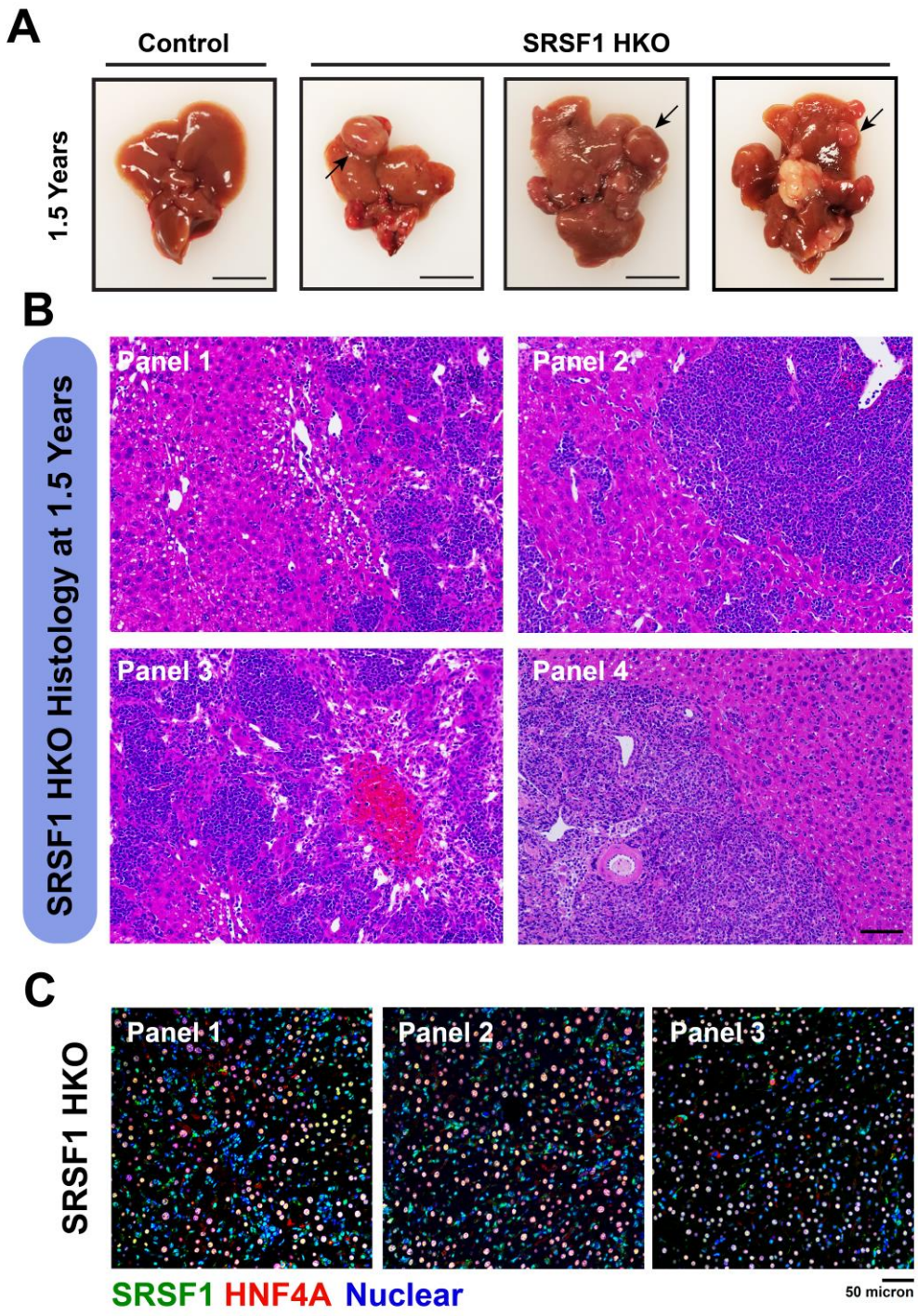


Figure 2.11: ***Characterization of 1.5-year-old SRSF1 HKO mice***

(A) Liver tissue from SRSF1 HKO develop growths on the liver (black arrow). (B) Histology of the tissue reveals severe infiltration of inflammatory cells into the tissue. (C) IF imaging shows that the tissue remains populated with SRSF1 expressing hepatocytes.

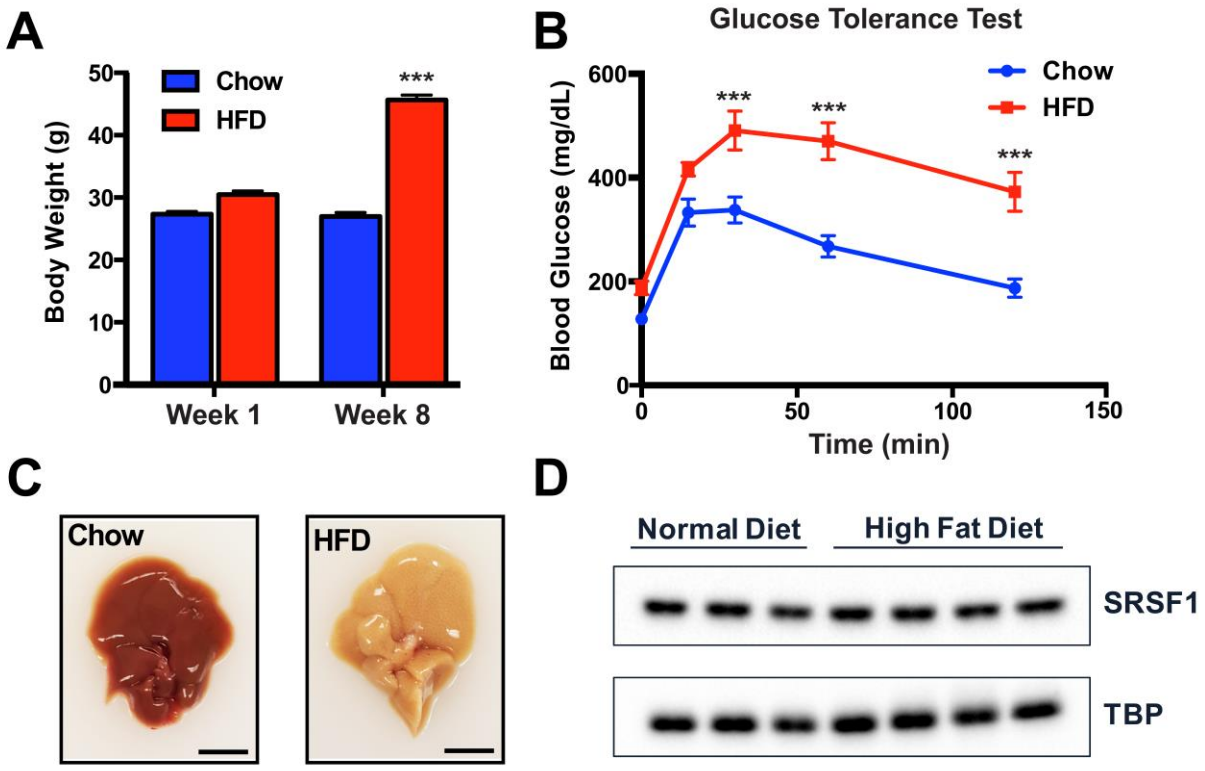


Figure 2.12: ***SRSF1 levels are not Regulated by Hepatic Lipid Accumulation***

A hepatic steatosis model was developed by feeding mice a high-fat diet. (A) After 8 weeks of feeding, mice were significantly heavier compared to controls. (B) High-fat diet mice have severe insulin resistance and (C) there is a striking accumulation of lipid within the tissue. (D) Western blot analysis on isolated hepatocytes from high-fat diet fed mice show no significant change in SRSF1 abundance.

Chapter 3: Transcriptome Changes in SRSF1 HKO Mice Model

3.1 Introduction

Knocking out SRSF1 in hepatocytes results in severe accumulation of lipids within the tissue and widespread liver damage. SRSF1, originally discovered as a splicing factor, is known to regulate gene expression post-transcriptionally. Because of this activity, it was expected the pathogenesis observed is due to defects in the transcriptome. To gain better insight into the misregulation occurring within the transcriptome, RNA-Seq analysis was performed on purified hepatocytes. This approach allows for unbiased view of the transcriptome changes occurring in the SRSF1 HKO mice hepatocytes. Specifically, it provides information regarding both gene expression and splicing differences occurring in the mice model. Having a complete representation of these changes will help in elucidating the mechanisms causing immediate death of hepatocytes deficient of SRSF1. RNA-Seq was performed at 10 days and 5 weeks of age to allow for correlation of the defects in the transcriptome with the pathogenesis seen from the histological analysis.

3.2 Material and Methods

3.2.1 Isolation of RNA, Library Preparation and RNA-Seq

Total RNA was prepared from frozen hepatocyte pellets using the RNeasy tissue mini-kit (Qiagen) following manufacturers protocol. Downstream RNA quality was assessed using an Agilent Bioanalyzer and quantified using a Qubit Fluorometer at the Roy J. Carver Biotechnology Center, UIUC. RNA-seq libraries were prepared and 100-bp paired-end Illumina sequencing was performed on a HiSeq 4000 at the High Throughput Sequencing and Genotyping Unit, UIUC.

3.2.2 RNA-Seq Analysis Workflow

RNA-Seq reads were processed for quality and read length filters using Trimmomatic (version 0.38) (Bolger, Lohse, and Usadel 2014). For differential gene expression analysis, RNA-Seq reads were psuedoaligned using Kallisto (version 0.44.0) (Bray et al. 2016) **(Figure 3.1)**. Transcript abundances were converted to gene abundances using tximport (version 1.11.7) (Soneson, Love, and Robinson 2015). With gene abundance tables, differential gene expression analysis was performed using DESeq2 (version 1.23.10) (Love, Huber, and Anders 2014). Differentially expressed genes were defined as having a $|Log2(FoldChange)| > 1$ and $FDR < 0.05$. For differential splicing analysis, RNA-Seq reads were mapped on the mouse genome (mm10) available on Gencode using STAR (version 2.4.2a) (Dobin et al. 2013). Alignment files were then used to perform differential splicing analysis using rMATS (version 3.2.5) and significant events were identified using imposed cutoffs ($FDR < 0.10$, junction read counts ≥ 10, and deltaPSI $\geq 15\%$) (Shen et al. 2014). Gene ontology analysis was performed using gProfiler and Enrichr (Raudvere et al. 2019; Kuleshov et al. 2016). Filtering and processing of data was performed using custom Python and R Scripts.

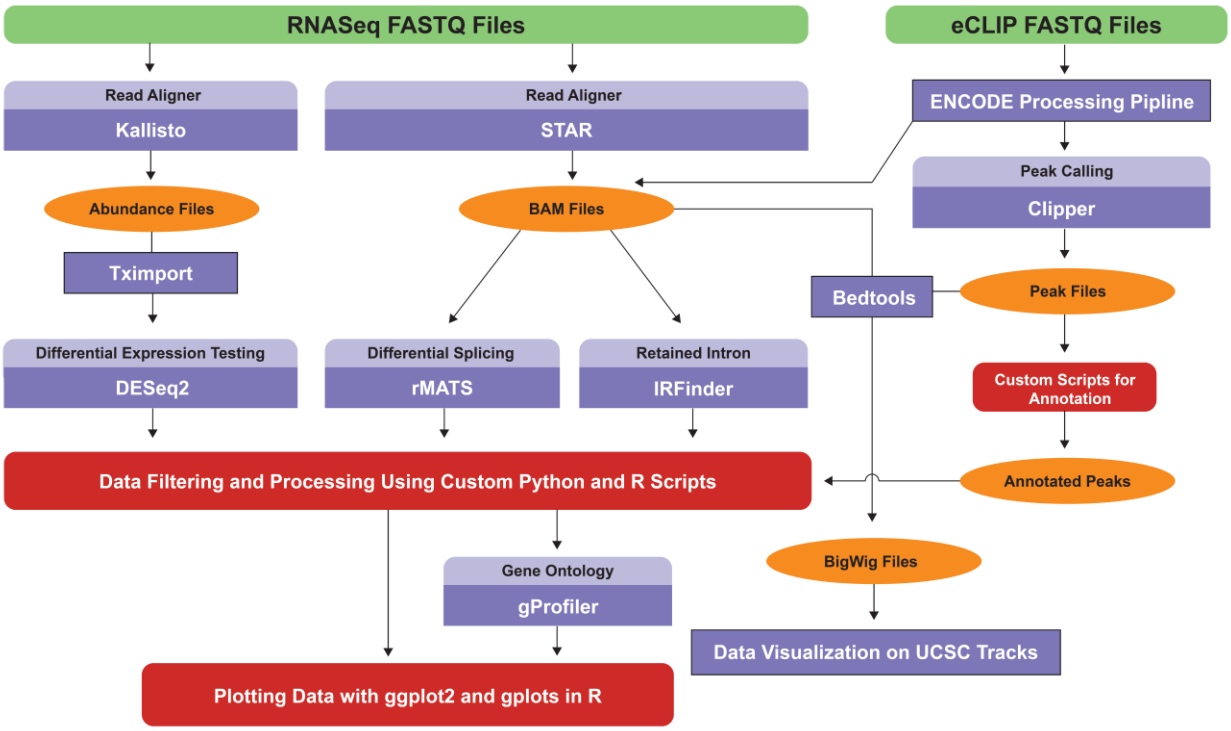


Figure 3.1: ***Transcriptome Data Analysis Pipeline***

A flowchart of the workflow and packages used to analyze transcriptome data.

3.3 Results

3.3.1 Early and Late Gene Expression Changes in SRSF1 HKO Hepatocytes

RNA-seq was performed on hepatocytes isolated from control and SRSF1 HKO at post-natal day 10 and 5-week timepoints to understand the early and late changes, respectively. Differential gene expression analysis reveals 1,825 and 1,902 genes changing in abundance at early and late stages, respectively **(Figure 3.2)**. Furthermore, a larger fraction of differentially expressed genes (DEGs) are increasing in expression. Specifically, ~63% of DEGs are upregulated at the early timepoint whereas ~79% are upregulated at the late timepoints. DEGs overlapping between the early and late phases encompass 544 genes which corresponds to one-third of either gene set. In summary, loss of SRSF1 significantly impacts the hepatocyte transcriptome at both early and late stages of development. Furthermore, a large fraction of these changes persists through development as the hepatocytes struggle to resolve the irregularities resulting from SRSF1 deficiency.

An interesting observation regarding the distribution of the overlapping gene set is that the fold change in abundance at the late phase tend be greater than the fold change during the early phase. This is apparent from the distribution of Log2(Fold Change) values **(Figure 3.4)**. The Pearson correlation of the linear fit for the early and late overlap gene set is strong with a value of 0.647 and a 1:1 correlation with a slope of 0.938. However, the linear fit is shifted vertically up with an intercept of 0.988. This signifies that the changes in abundance are greater in the late phase when compared to the abundance change estimated at the early timepoint. This can further be demonstrated by the cumulative frequency plot of the differences in Log2(Fold Change) values between late

and early timepoints for the overlapping gene set **(Figure 3.5)**. In comparison to the control set which consists of all expressed genes, the overlapping set is shifted to the right with a median difference of 0.648 which corresponds to a fold change of 1.56. Therefore, the genes that are upregulated in the early timepoint are further potentiated in expression at the later timepoint.

3.3.2 Gene Ontology Analysis for Differentially Expressed Genes in SRSF1 HKO

It is evident that substantial alterations to the transcriptome are triggered in response to SRSF1 ablation. It can be hypothesized that these changes are induced to resolve or compensate the irregularities caused by the deficiency of SRSF1 activity. To gain an understanding into what biological processes the DEGs are involved with, a gene ontology analysis was performed **(Figure 3.3)**. The analysis was divided by the genes changing in expression at each timepoint and the direction of the change in abundance. Gene ontology analysis was performed using Enrichr and the results are represented in a bar plot with each GO term plotted with its Combined Score. Larger combined score values correspond to stronger enrichment of the GO term.

Upregulated genes at the early and late timepoints enriched for terms involved in separate processes. For genes upregulated during the early SRSF1 HKO timepoint enriched for processes involved in inflammation with terms such as ‘positive regulation of chemotaxis’ and ‘chemokine-mediated signaling pathway’. This is not unexpected since knockout of SRSF1 results in immediate death of the cell as seen previously. It is known that widespread cell death triggers an inflammatory response to mediate recruitment of cells for repair of the tissue (Rock and Kono 2008). However, genes upregulated at the late timepoint enrich for processes involved in mitosis. This result agrees with the

phenotypic characterization of this model. At the 5-week timepoint, the SRSF1 HKO are still in the process of repopulating the tissue with hepatocytes that have shutdown Cre expression and can express SRSF1. As the tissue continues to repopulate, the inflammation subdues and proliferation increases. Biological processes that are shared at both timepoints are involved in wound healing and repair.

Downregulated genes enrich for processes involved in metabolism at both timepoints. However, each timepoint is downregulated for a specific aspect of metabolism. Genes downregulated at the early timepoint are involved in sterol transport and fatty acid oxidation pathways. Whereas genes downregulated at the late stage encompass sulfur amino acid, ornithine, and urea metabolism. These differences in metabolic processes are most likely driven by differences in metabolism of normal livers as it matures.

3.3.3 Inflammation and Wound Healing Signatures in SRSF1 HKO Hepatocytes

Histological characterization of SRFS1 HKO mice livers across developmental stages shows significant damage of the tissue. Earlier stages are predominated by inflammation and cell damage. However, with time the inflammation subdues, and fibrosis occurs as scar tissue develops. To see if these findings are corroborated by gene expression profiles, levels of various inflammatory and fibrosis markers were examined from RNA-seq **(Figure 3.6)**. In agreement with phenotypic characterization, inflammatory genes are primarily upregulated in the 10-day timepoint while fibrotic signature is activated in the 1-month timepoint.

3.3.4 Differential Splicing of Exons in SRSF1 HKO Hepatocytes

Due to SRSF1's role as a splicing factor, it is expected that loss of activity will result in defects in splicing of exons. To further gain insight into the splicing misregulation, differential splicing analysis was performed on the RNA-seq data using the well-established rMATS package. Performing this analysis on the 10-day and 1-month timepoint datasets revealed 852 and 765 differentially spliced events, respectively **(Figure 3.7 A)**. The overlap between these two sets consist of 216 splicing events. The splicing events are further sub-divided by the type of alternative splicing event which includes skipped exon (SE), mutually-exclusive exon (MXE), retained intron (RI), alternative 5' splice site (A5SS), and alternative 3' splice site (A3SS). At both timepoints, the SE event type had the largest representation of differentially spliced exons **(Figure 3.7 B)**. Furthermore, loss of SRSF1 resulted in over two-thirds of exons with decreased inclusions **(Figure 3.9 A)**. This is expected as SRSF1 is known to promote the inclusion of exons. The differentially spliced events which were detected at both timepoints had very similar distributions. This can be seen in the distribution of ΔPSI values of the overlapping events. The correlation of the linear fit is strong with a slope of 0.856. Furthermore, the distribution of the difference of the ΔPSI (ΔΔPSI) are mostly centered around zero **(Figure 3.9 B)**.

3.3.5 Gene Ontology Analysis for Differentially Spliced Exons in SRSF1 HKO

In comparison to the gene ontology analysis performed on the differential gene expression dataset, the analysis for differential splicing showed weaker enrichment scores. Genes with differentially spliced exons at the 10-day timepoint enriched for processes involved in DNA damage while the 1-month timepoint enriched for processes

involved in nuclear protein localization and RNA destabilization **(Figure 3.10 A)**. Genes that shared differentially spliced exons at both timepoints enriched for terms involved in wound healing and regulation of signaling cascades **(Figure 3.10 B)**.

3.3.6 Overlap of Differential Expression and Splicing in SRSF1 HKO

To examine if the changes in gene expression and splicing are dependent processes, overlap of these two datasets at each timepoints was determined. For example, if splicing has a significant effect on steady state gene abundance, it would be expected that a significant overlap of the two sets be present. Overlap of gene expression and splicing showed a total of 72 and 78 genes at the 10-day and 1-month, timepoints **(Figure 3.11)**. Furthermore, the overlap p-values at both timepoints showed that the overlap was not significant (overlap p-value < 0.477 and < 0.275 at 10-day and 1-month, respectively). Therefore, this suggests that changes in splicing and gene expression are primarily independent.

3.4 Conclusions

RNA-Seq analysis of SRSF1 HKO mice at both ten-day and 1-month timepoints provide rich information regarding changes in gene abundances and exon usage occurring in the model. It is apparent that loss of SRSF1 activity in hepatocytes results in substantial changes to the SRSF1 HKO transcriptome. It should be noted, however, that phenotypic characterization showed that by post-natal day 10, significant liver damage and hepatic cell death has already occurred. Therefore, the changes that are detected in the transcriptome are most likely a result of secondary effects of SRFS1 ablation. Deficiency of hepatic activity triggers immediate regenerative response as early as 10 days. This is further corroborated with the gene expression signatures seen in the 10-day RNA-seq. Namely, an enrichment for processes involved in inflammation is seen in the upregulated gene set. This response is expected as widespread cell death will inevitably trigger an inflammatory response (Rock and Kono 2008). However, as the liver continues to regenerate it spontaneously repopulates with SRSF1-expression hepatocytes as it gains the ability to suppress Cre expression. At later timepoints, the inflammation reduces, and further expansion of the SRSF1-expressing hepatocytes predominates to compensate for the required hepatic functionality. Indeed, by the later 1-month timepoint, a strong enrichment for genes involved in mitosis and cell cycle processes are enriched **(Figure 3.3)**.

As expected, due to the widespread damage and loss of hepatocyte functions, a strong decrease in metabolic capacity is seen in the SRSF1 HKO. Interestingly, each timepoint exhibits a downregulation a specific aspect of metabolism. Specifically, early timepoint shows downregulation of genes involved in fatty acid metabolism whereas late

timepoint shows downregulation of amino acid, ornithine, and urea cycle metabolism. A possible explanation for this difference is that it is driven primarily by the difference of the metabolic gene profile during liver development. As the liver matures, there is activation of mature metabolic functions including amino acid and urea cycle metabolism. Therefore, metabolic genes in general are downregulated in SRSF1 HKO due to the widespread damage response.

Overall, the RNA-seq data at these two timepoints further support the pathology observed in the SRSF1 HKO mice. These mice exhibit significant inflammation at early stages which decreases with time as the tissue repopulates with SRSF1-expressing hepatocytes. At later stages, development of scar tissue is observed with the development of fibrosis. Indeed, these signatures are also seen from gene expression analysis at the early and late timepoints. Specifically, inflammatory signature is activated at the 10-day timepoint with fibrotic signature primarily upregulated at the 1-month timepoint **(Figure 3.6)**. Due to the immediate damage response, to gain better insight into the mechanism of pathogenesis, a model allowing for improved temporal control of SRSF1 ablation will provide opportunity to isolate hepatocytes before secondary changes occur.

3.5 Chapter Figures

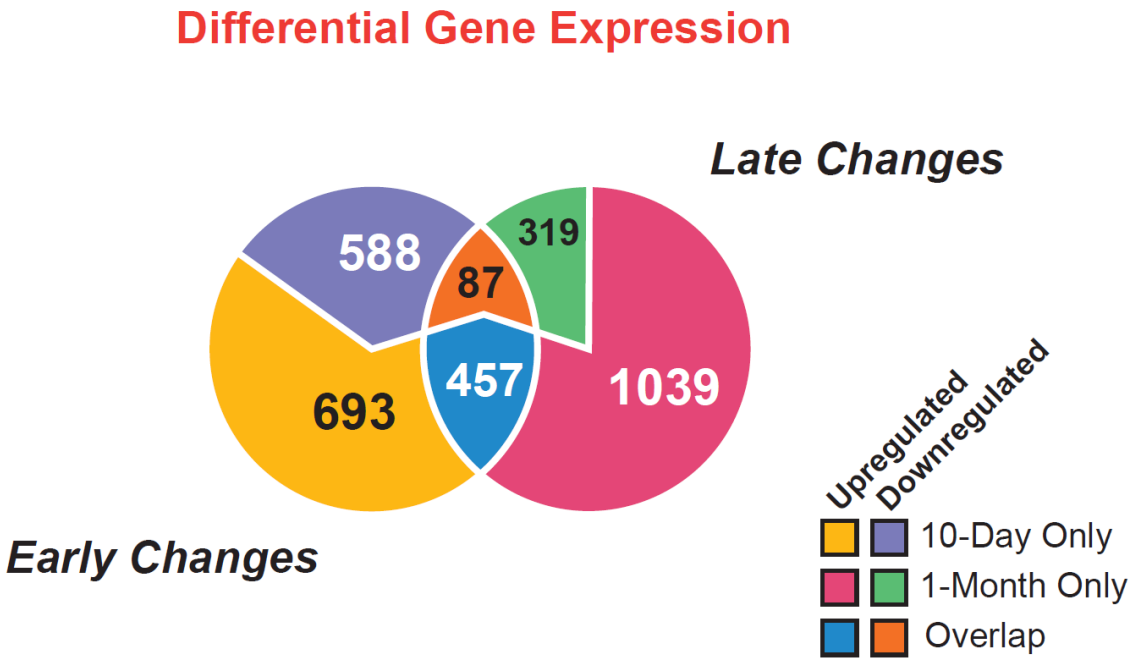


Figure 3.2: ***Breakdown of Early and Late Gene Expression Changes in SRSF1 HKO***

A Venn diagram summarizing the number of genes changing in expression in hepatocytes of SRSF1 HKO at both early and late stages of development. A total of 1825 and 1902 genes are changing in expression (Log2 Fold Change > 2 and FDR < 0.1) at early and late stages, respectively. At both stages, a greater proportion of genes are being upregulated.

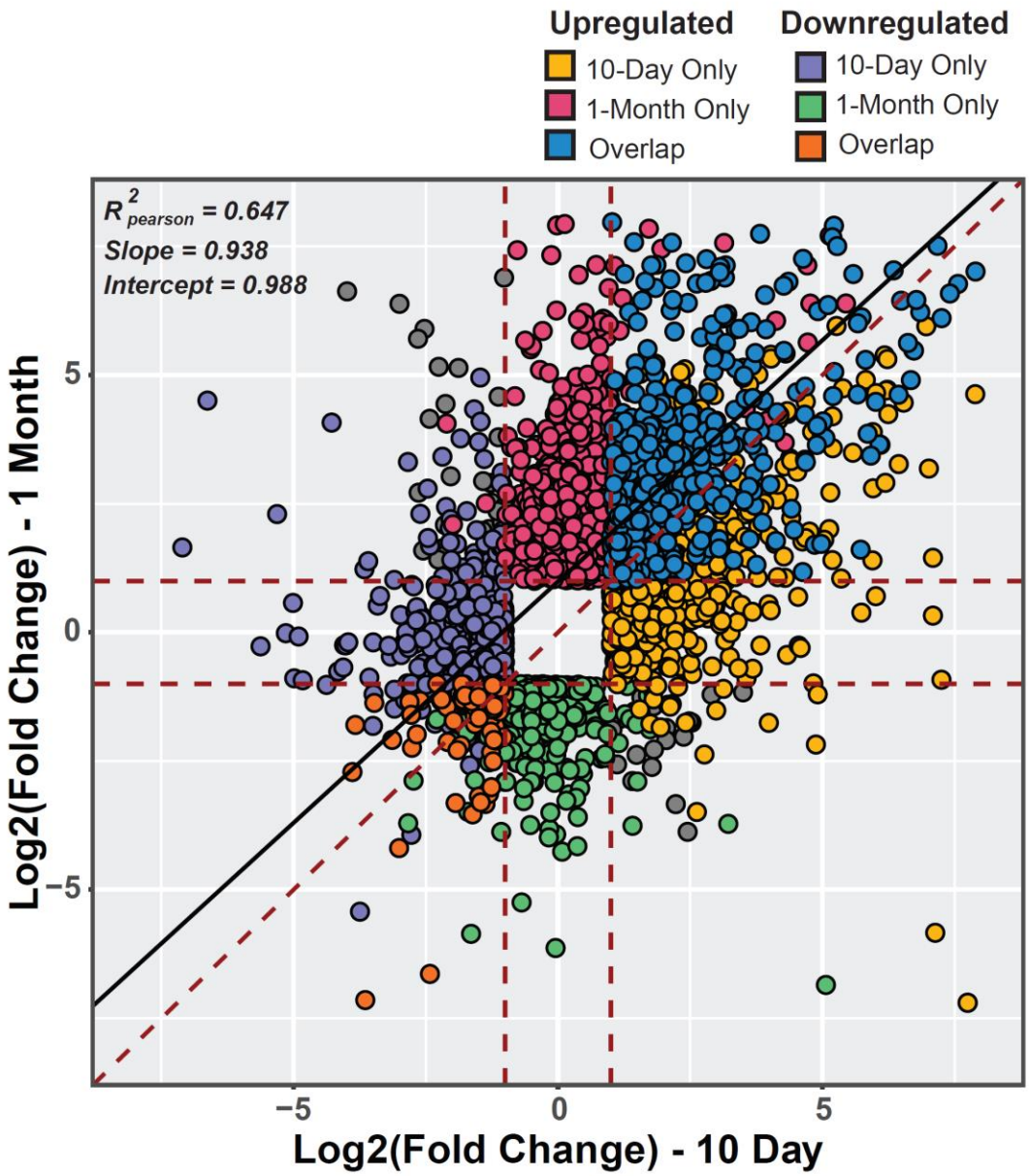


Figure 3.3: ***Scatter Plot of Expression Fold Changes of Early and Late Stages***

Scatter plot of gene expression fold change values at both 10 days and 1-month timepoints. Significant number of genes expression changes are specific to each timepoint. A strong positive correlation is seen for genes that are found to be changing at both timepoints.

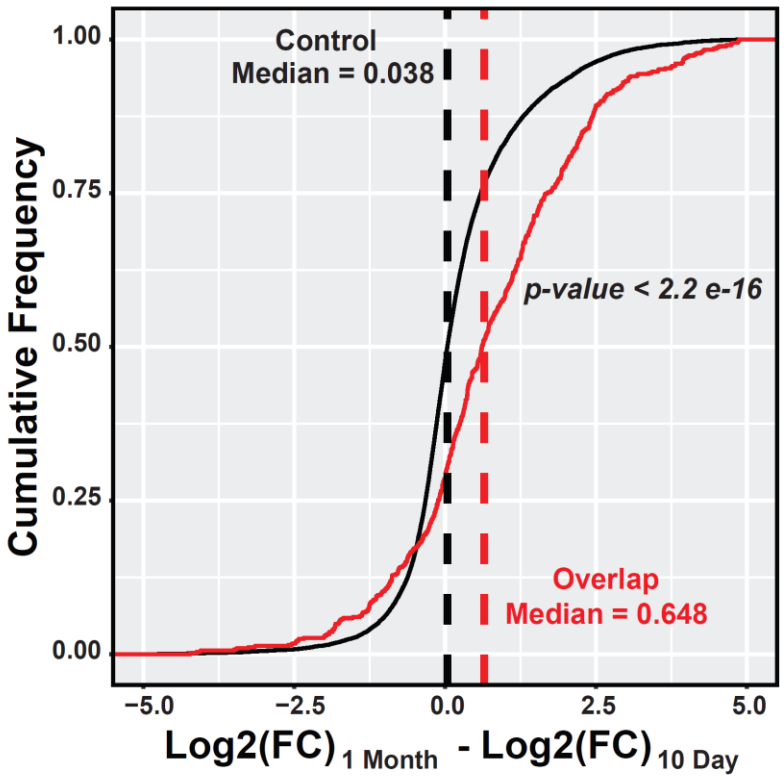


Figure 3.4: ***Cumulative Plot of Fold Change Differences for Differentially Expressed Genes in Late and Early Timepoints***

Cumulative plot of fold change differences for genes differentially expressed in early and late timepoints (red line). The Control plot (black line) corresponds to fold change differences for all expressed genes between both timepoints. The cumulative plot of the overlap differences is right shifted, signifying that for genes found to be differentially expressed at both timepoints, the induction of expression is greater at the late timepoint.

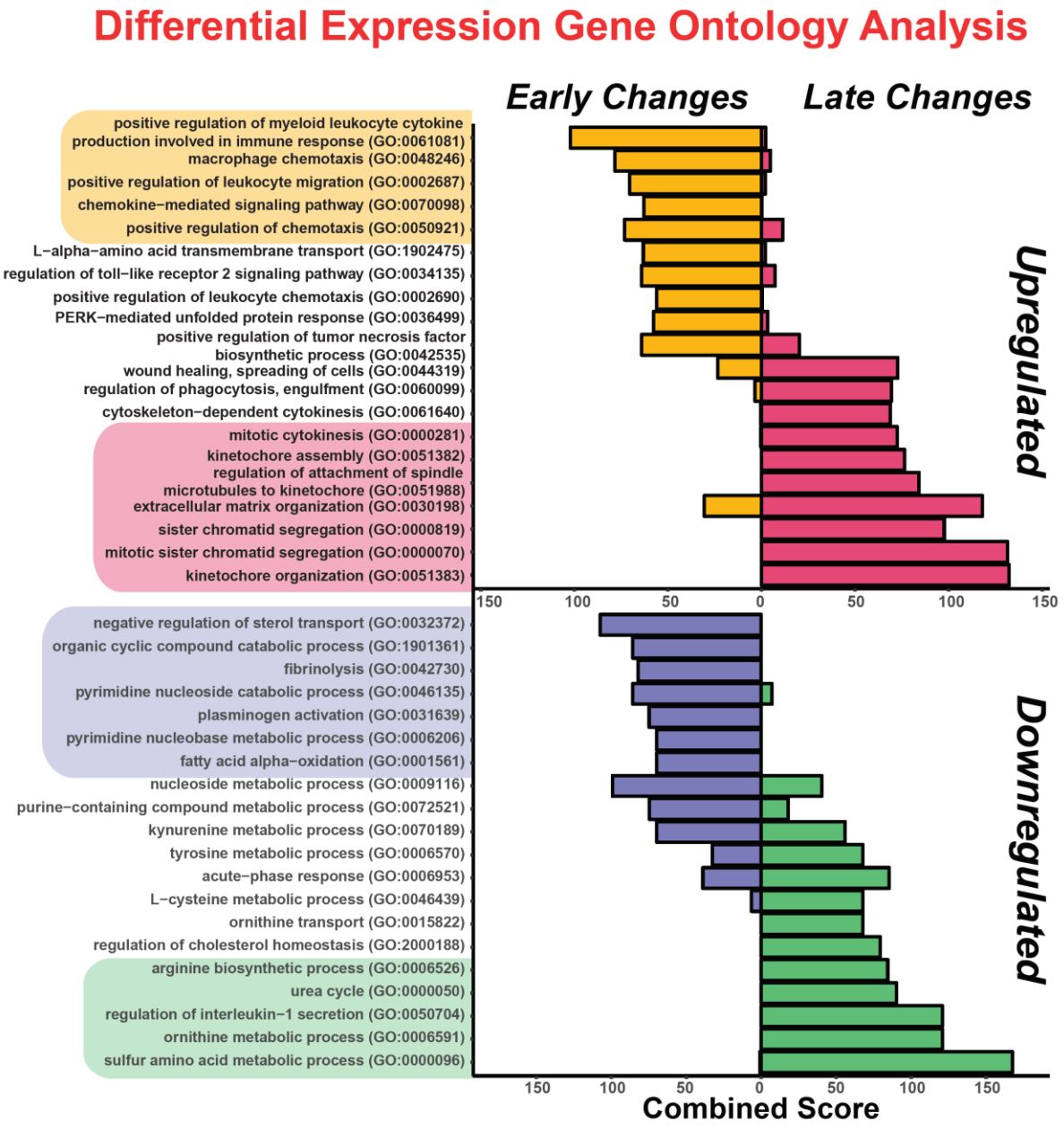


Figure 3.5: ***Gene Ontology Analysis of Differentially Expressed Genes in SRSF1 HKO***

Gene ontology analysis of differentially expressed genes during early and late timepoints. Genes upregulated during early timepoints enrich for inflammatory processes while genes activated at late timepoints enrich for cell cycle and mitotic processes. During both timepoints, different aspects of metabolic processes are downregulated.

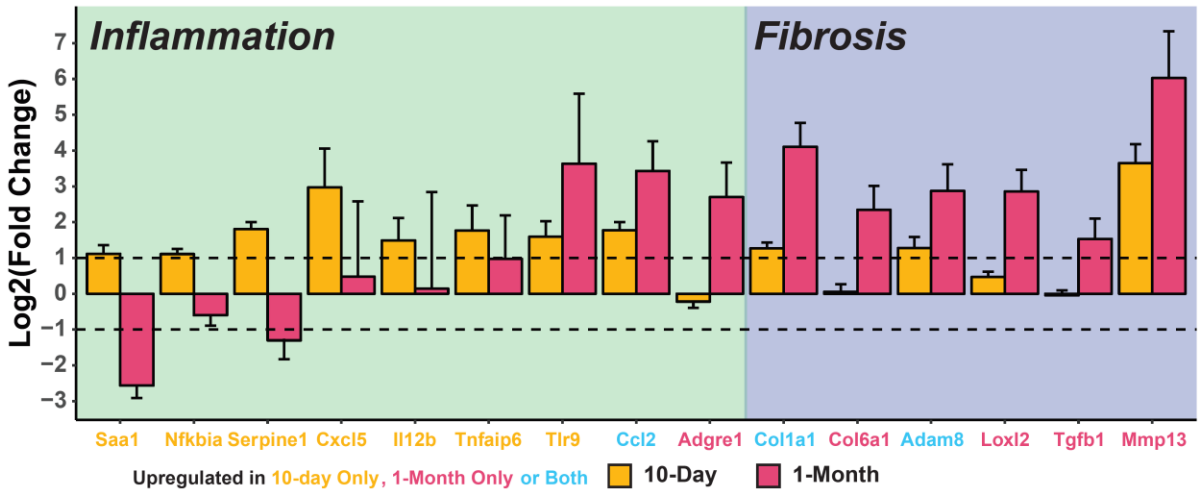
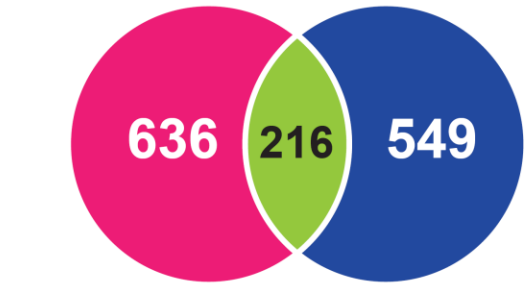


Figure 3.6: ***SRSF1 HKO Hepatocytes Transition from an Inflammatory to Fibrotic Response with Development***

Plot of fold change values for genes involved in inflammation and fibrosis at both early (yellow bars) and late (red bars) timepoints. At early timepoints, a larger fraction of genes involved in inflammation are induced. While genes involved in fibrosis are more strongly induced at the 1-month timepoint. This data is in support of the histological observations.

A



B

Alternative Splicing Type		10 Day	1 Month	Overlap
SE		199 434	165 407	36 139
MXE		20 24	14 20	4 7
RI		24 36	31 33	2 1
A5SS		8 43	11 40	1 12
A3SS		18 46	15 29	3 11
ΔPSI	Increased	269 (31.6%)	236 (30.8%)	46 (21.3%)
	Decreased	583 (68.4%)	529 (69.2%)	170 (78.7%)
Total Number of Events		852	765	216

Figure 3.7: ***Differential Splicing during Early and Late Stages of SRSF1 HKO***

(A) Venn diagram depicting the total number of exonic events with significant differential splicing (ΔPSI > 0.15, FDR < 0.10) between control and SRSF1 HKO mice at 10 days and 1-month timepoints. A percent overlap of about 20.2% and 22.0% was found between 10 days and 1-month SRSF1 HKO, respectively. (B) Breakdown of the number of differentially spliced exons by event type. A greater proportion of events showed decreased inclusion at both timepoints.

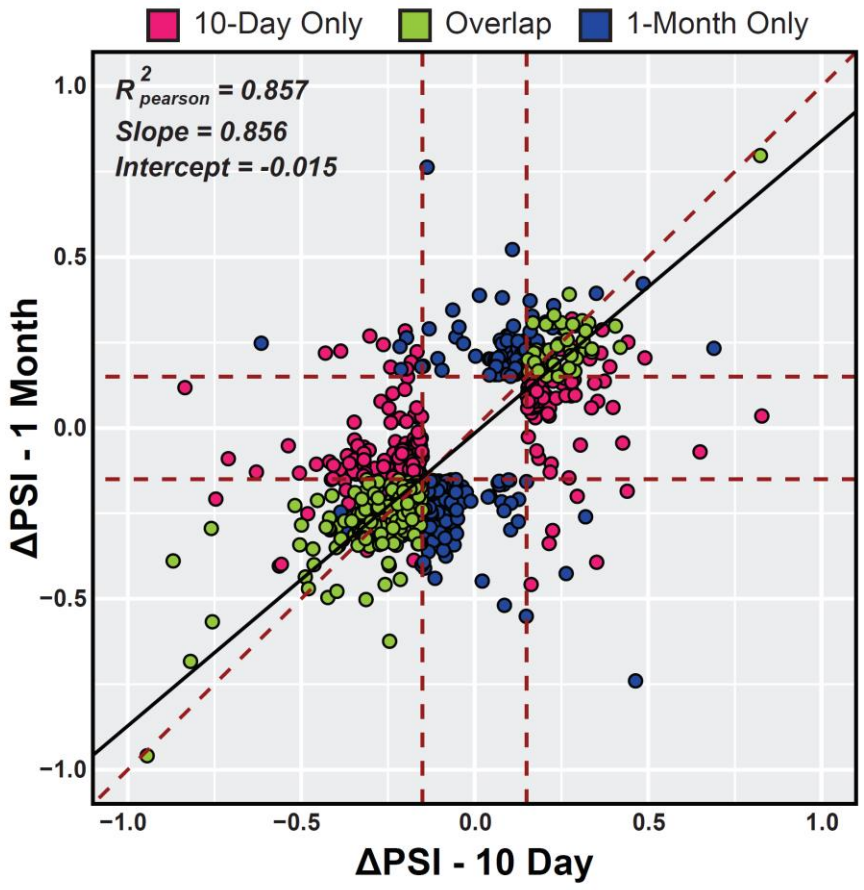


Figure 3.8: ***Scatter Plot of ΔPSI of Exons at Early and Late Stages of SRSF1 HKO***

Scatter plot of ΔPSI values for all significantly changing splicing events in either 10 days or 1-month SRSF1 HKO timepoints. Red points depict events which are specifically found to be changing at early stages while blue points are events which are late stage specific. Green points are events which are differentially spliced in both stages. Overlapping events have a strong 1:1 correlation.

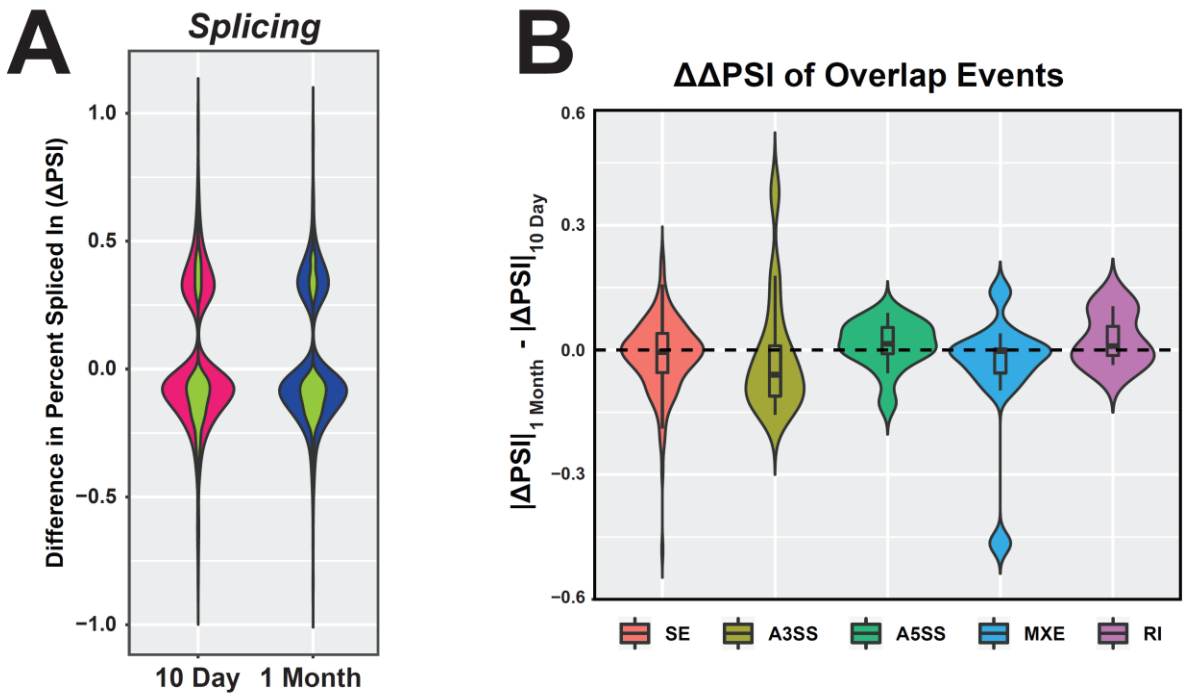


Figure 3.9: ***Distribution of ΔPSI values at Early and Late SRSF1 HKO Timepoints***

(A) Violin plot depicting distribution of ΔPSI values for all significant differentially spliced events at each timepoint. Green highlighted distribution signifies the differential exon events that are present in both timepoints. The distributions are not significantly different, and both have greater proportion of exons being excluded with SRSF1 knockout. (B) Violin plot depicting distribution of ΔPSI differences (ΔΔPSI) of overlapping events in early and late stages. All events type distributions are primarily centered at zero and thus showing high 1:1 correlation of overlapping events.

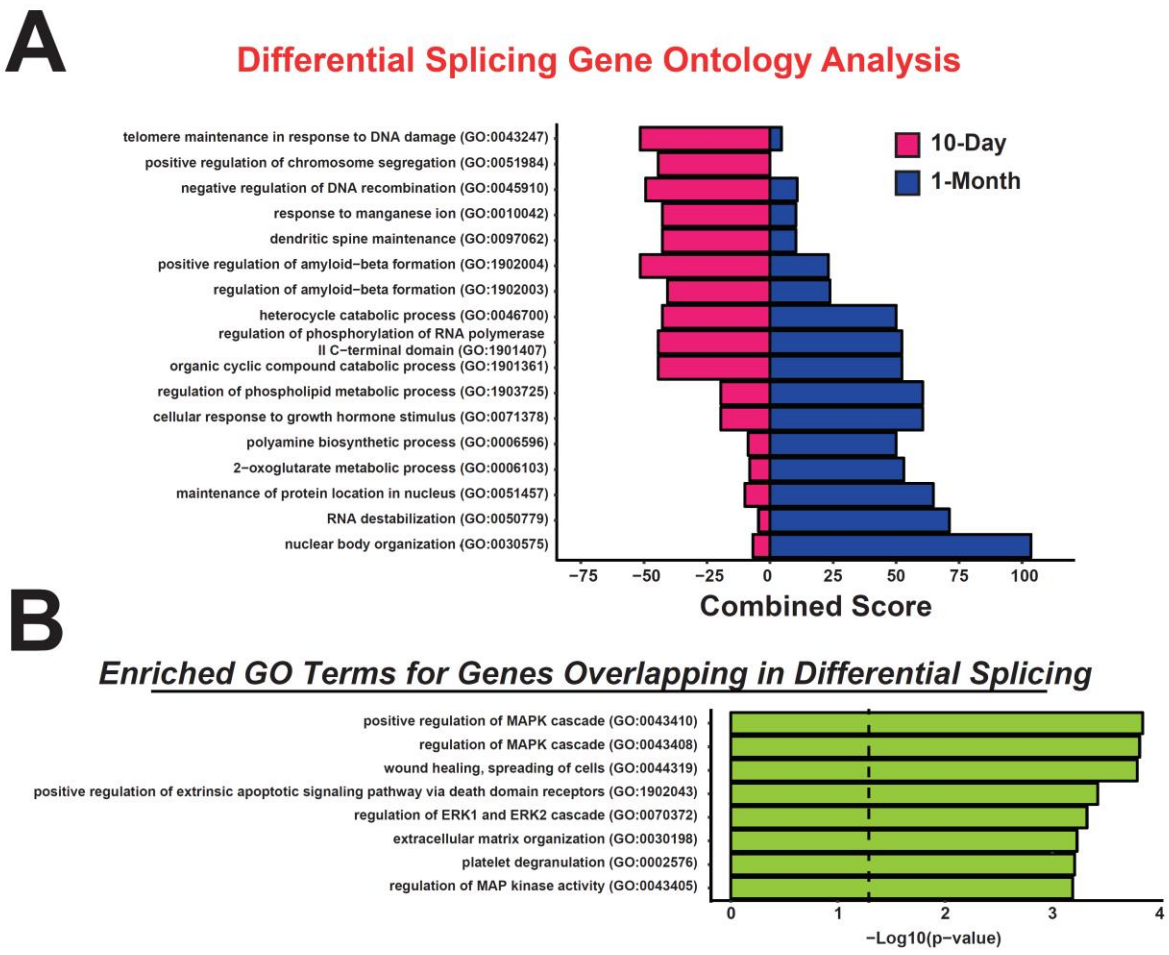


Figure 3.10: ***Gene Ontology Analysis of Genes with Differential Splicing at Early and Late Timepoints in SRSF1 HKO Mice***

(A) Gene ontology analysis of genes with differential splicing events at both early and late timepoints in SRSF1 HKO hepatocytes. Early phase genes with differential splicing are enriched for processes involved in DNA damage response and repair. While in the late phase, genes with differential splicing enrich for processes involved in nuclear protein organization and RNA destabilization. (B) GO terms for genes that are differentially spliced at both timepoints enrich for processes involved in wound healing and MAPK signalling.

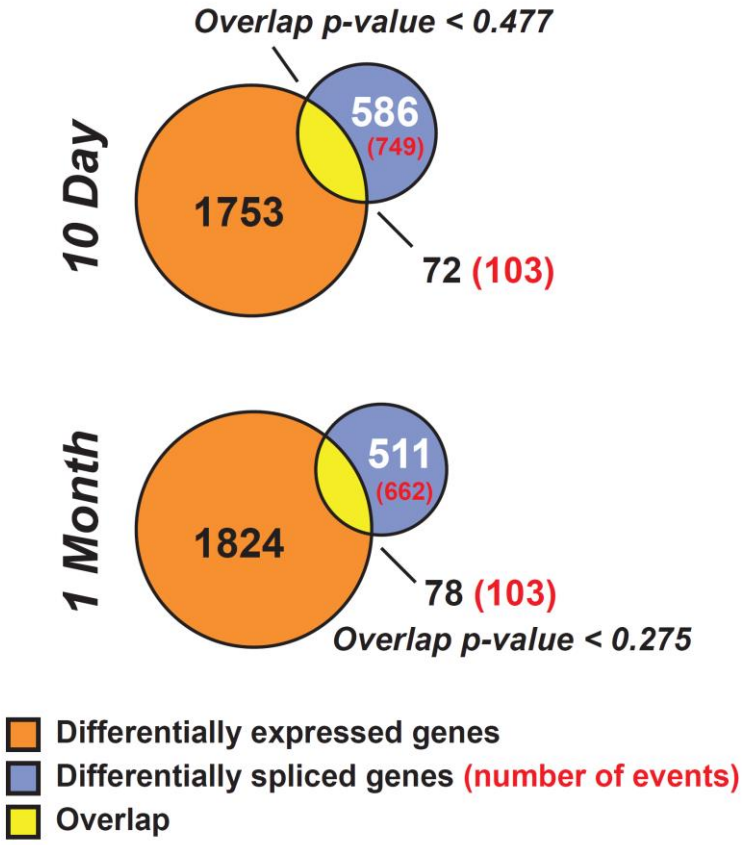


Figure 3.11: ***Genes with Differential Expression and Splicing in SRSF1 HKO for each Timepoint***

Venn diagrams depicting overlap of genes changing in expression and splicing for each timepoint. The overlap in each set is not significant (overlap p-value > 0.05) and thus providing evidence that these two forms of regulation are independent of each other.

Chapter 4: Elucidating Early Changes in Hepatocytes After Acute Ablation of SRSF1

4.1 Introduction

From Chapters 2 and 3, it is apparent that loss of SRSF1 within hepatocytes is not well tolerated and results in cell death. However, due to the high regenerative capacity of the liver, the tissue can subvert the knockout of SRSF1. This is most likely achieved through mechanisms such as transdifferentiation of BECs or asymmetric proliferation of hepatocytes resulting in aneuploid cells deficient of the Cre transgene. Repopulation of the liver with SRSF1 expressing hepatocytes is observed as early as eight days after birth. This aspect of the SRSF1 HKO model poses difficulty in elucidating the mechanism of pathogenesis, namely, the significant accumulation of lipids within the tissue and subsequent cell death. To circumvent this issue, an acute hepatocyte-specific SRSF1 knockout model (acSRSF1 HKO) was developed. This was achieved by using adeno-associated viral vectors expressing the Cre recombinase driven by a hepatocyte-specific promoter. This ensures that knockout of SRSF1 occurs only within hepatocytes. The acSRSF1 HKO allows for improved temporal control and studying of early changes occurring in the tissue upon loss of SRSF1. The acSRSF1 HKO allows for the isolation of SRSF1 deficient hepatocytes prior to any development of severe pathology. This model provides the ability to tease out the chain of events triggered by the loss of SRSF1 which results in the eventual death of the cell.

4.2 Material and Methods

The following sections provides details of the materials and methodology utilized for the studies performed in this chapter. For details on additional methodology refer to *Section 9.1: Supplementary Protocols.*

4.2.1 Generating Acute Hepatocyte-Specific SRSF1 Knock-out Mice Model

Mice with floxed alleles for SRSF1 were injected with adeno-associated viral vectors (VectorBio Labs) either expressing GFP (Controls) or the Cre recombinase (acSRSF1 HKO) from the thyroxine binding globulin (TBG) promoter which is hepatocyte specific. Mice were injected at 8 weeks of age with a titer of 5 x 10^{11} gc via the tail vein. The mice were housed on a standard 12-hour-light/dark cycle and fed a normal chow diet (2918 Envigo Teklad). The mice were fasted for 6 hours prior to harvesting tissues for subsequent analysis. Tissue harvesting was performed at 2- and 4-weeks post-injection. Hepatocytes were isolated only after 2 weeks. National Institutes of Health (NIH) guidelines for use and care of laboratory animals was followed and all the experiments were approved by the Institutional Animal Care and Use Committee at the University of Illinois at Urbana-Champaign.

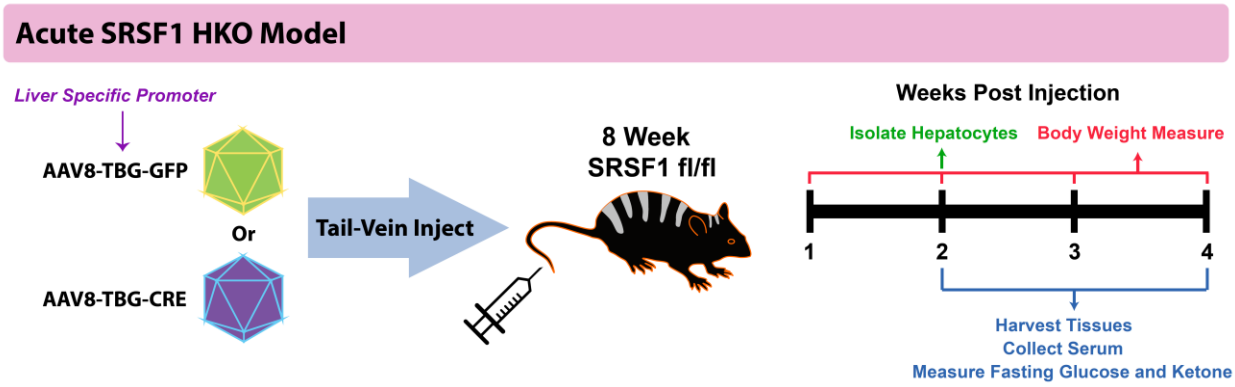


Figure 4.1: ***Schematic of Acute SRSF1 HKO Model***

4.2.2 Lipid Isolation from Liver Tissue

Lipids were extracted from approximately 100 mg of liver tissue using the Folch Method (Folch, Lees, and Sloane 1957). Briefly, the tissue was homogenized in 1 mL of 2:1 choroform:methanol mixture. The homogenized mixture was incubated overnight at room temperature to allow for complete extraction. The mixture was then centrifuged at 10,000 x g for 10 minutes to pellet any debris. The supernatant was transferred to a fresh tube. The interface was washed with 1 X PBS solution to remove any additional salt. The mixture was centrifuged at 4,000 x g and aqueous layer was removed. The organic phase was dried under nitrogen stream and then reconstituted in 300 µL of ethanol. The reconstituted mixture was used with Infinity Kit to determine concentrations of triglycerides and cholesterol.

4.2.3 Serum Chemistry Analysis

Serum chemistry measurements were performed in both control and acSRSF1 HKO mice at 2- and 4- weeks post viral injections. Mice were fasted for 6 hours (noon to 6 p.m.) before collecting blood. Blood was collected as specified in *Section 2.2.2*, followed by flash freezing of the serum in liquid nitrogen. Frozen serum samples were submitted to the Mouse Metabolic Phenotyping Core at the University of Cincinnati. A clinical chemistry analyzer was used to measure concentration of the following serum components, triglycerides, cholesterol, phospholipids, non-esterified fatty acids, glucose, ketones, ALT and AST.

4.3 Results

4.3.1 Acute Hepatocyte-Specific Knockout of SRSF1 Results in Severe Steatosis

Using the AAV8-TBG-iCre vector, robust hepatocyte-specific knockout of SRSF1 was achieved. By 2-weeks post-viral transduction, near complete loss of SRSF1 expression can be seen by western blot analysis **(Figure 4.2 B)**. At this timepoint, gross inspection of the liver shows mild changes in appearance corresponding to mild steatosis. However, by 4-weeks, the acSRSF1 HKO livers are notably yellow in color throughout the whole liver signifying severe steatosis **(Figure 4.2 A)**. Serum collected at this timepoint is also strikingly golden in color which is most likely due to increased bilirubin levels **(Figure 4.2 C)**. Furthermore, weekly assessment of body weight changes in the acSRSF1 HKO model shows that these mice maintain their body weight until 3-weeks post viral transduction at which point the mice begin to drastically lose weight. By 4-weeks the acSRSF1 HKO mice lose on average about 15% of their body weight **(Figure 4.3)**.

4.3.2 Eventual Repopulation of Liver with SRSF1 Expressing Hepatocytes

Interestingly, SRSF1 western blot analysis at the 4-week timepoint shows increased levels of SRSF1 in comparison to the 2-week timepoint **(Figure 4.2 B)**. To determine if these increased levels are due to infiltration of the tissue with inflammatory cells or repopulation with SRSF1-expressing hepatocytes, immunofluorescent co-staining for SRSF1 and HNF4α was performed on tissue sections **(Figure 4.4 A)**. The fluorescent images from acSRSF1 HKO reveal that there is repopulation of the tissue with SRSF1-expressing hepatocytes. This expansion also appears to originate from the portal triad region. To assess for apoptosis in the acSRSF1 HKO model, TUNEL staining was performed which were negative at both timepoints **(Figure 4.4 B)**.

4.3.3 Histological Analysis Reveals Hepatocyte Necrosis in acSRSF1 HKO

To gain insight into the pathological changes occurring at the level of tissue architecture, histological staining of tissue sections was performed. H&E, Oil Red O and Sirius Red staining were performed on tissues collected from control and acSRSF1 HKO mice at the 2- and 4-week timepoints. At the 2-week timepoint, the tissue is beginning to show lipid accumulation with microsteatosis occurring near central vein hepatocytes or Zone 1. At this timepoint minimal damage is observed in the tissue architecture with no development of fibrosis as seen by the absence of staining in Sirius Red. However, by 4 weeks, Oil Red O staining shows significant limpid accumulation with macrosteatosis present within all zones of the tissue. Furthermore, H&E staining shows ballooned hepatocytes which suggest necrotic cell death. Similar to the 2-week timepoint, no scarring of the tissue is observed.

4.3.4 Loss of SRSF1 Leads to Decreased Levels of Hepatic Cholesterol

Gross inspection of the tissue shows that the acSRSF1 HKO model develop severe steatosis. However, it is not apparent whether this accumulation of lipids is due to triglycerides or cholesterol. To answer this question, total lipids was extracted from the tissue using the Folch method followed by quantification of triglycerides and cholesterol levels normalized to the tissue mass. This assay was performed in controls and acSRFSF1 HKO at both 2- and 4-week timepoints. The quantification results of the hepatic triglycerides and cholesterol show that the composition of the accumulated lipid are triglycerides **(Figure 4.6)**. The triglycerides progressively increase in concentration with time which agrees with the gross appearance noted of the liver. Interestingly, hepatic cholesterol levels are diminished in acSRSF1 HKO.

4.3.5 AcSRSF1 HKO Experience Depletion of Adipose Stores

Characterization of the acSRSF1 HKO mice reveals striking lipid accumulation within the liver tissue and significant loss in total body weight. To gain additional insight into the state of both liver and adipose tissues in acSRSF1 HKO, liver and gonadal adipose tissues were weighed and normalized to total body weight at each timepoint **(Figure 4.7)**. As expected, liver weight normalized to body weight progressively increase as the tissue accumulates lipids. Interestingly, on the other hand, adipose tissue shows a striking decrease in normalized weights. At the 4-week timepoint, there is nearly complete loss of the gonadal adipose tissue stores.

4.3.6 Serum Profiling Reveals Loss of Hepatic Function in acSRFS1 HKO

The liver plays a major role in maintaining levels of multiple metabolites within the serum. Understanding how the serum is affected in the acSRSF1 HKO will provide better grasp of the deficiencies caused by loss of SRSF1 activity. To this end, serum profiling was performed on 6-hour fasted control and acSRSF1 HKO mice at each timepoint **(Figure 4.8)**. Overall, 2-week acSRSF1 HKO do not show any significant differences of any serum parameters. This provides support that significant damage and secondary effects have not yet occurred in this model. However, by 4 weeks, acSRSF1 HKO mice show significant elevation in ALT and AST levels signifying severe liver damage. With regards to fasting glucose levels, 4-week acSRSF1 HKO maintain a level around 65 mg/dL which is significantly decreased in comparison to controls. Furthermore, ketone bodies are also strikingly diminished at the 4-week timepoint. Finally, serum cholesterol and phospholipid levels show progressive decline in concentration. Triglycerides and NEFA levels do not exhibit significant differences in serum level at either timepoint.

4.4 Conclusions

Due to the immediate liver damage that is triggered in the SRSF1 HKO mice, the acSRSF1 HKO mice was developed. In this model, knockout of SRSF1 is induced by transduction of an AAV vector expressing the Cre transgene into SRSF1 floxed mice. This provides further temporal control of over the knockout of SRSF1. Furthermore, this knockout is hepatocyte-specific as the Cre transgene is driven by the hepatocyte-specific TBG promoter. The temporal control of SRSF1 knockout allows for isolation of hepatocytes before severe damage and secondary effects are triggered. However, before any molecular studies can be performed, a thorough phenotypic characterization of the model is necessitated. The first insight this model provides is whether the damage observed in SRSF1 HKO dependent on the developmental stage at which the knockout occurs. It will also reveal if knocking out SRSF1 in hepatocytes in adult livers phenocopy the SRSF1 HKO mice which experience knockout at post-natal day 6. Indeed, knocking out of SRSF1 in adult mice livers resulted in severe steatosis and liver damage as observed in the SRSF1 HKO mice **(Figure 4.2 and 4.5)**. Interestingly, however, the acSRSF1 HKO appear to develop the pathology at a slower rate as compared to the SRSF1 HKO model. For instance, in the SRSF1 HKO model, knockout is observed at post-natal day 6 with significant steatosis within 4 days at post-natal day 10. In the acSRSF1 HKO, knockout is observed at 2 weeks with significant steatosis and damage occurring at 4 weeks. It should be mentioned that intermediate timepoints were not collected for the acSRSF1 HKO, therefore, the exact timing at which pathology develops in acSRSF1 HKO mice cannot be determined. However, the extent of repopulation with SRSF1-expressing hepatocytes at 4 weeks is comparable to what is seen at the 10-day SRSF1 HKO mice **(Figure 4.4)**. Furthermore, the weights of acSRSF1 HKO mice are

stable until 4-weeks at which a significant drop is observed **(Figure 4.3)**. Taking this together, the development of pathology in the acSRSF1 HKO mice appear to be delayed in comparison to SRSF1 HKO. A possible explanation of this observation is due to the developmental timing at which the knockout occurs. In SRSF1 HKO, the knockout occurs at an early timepoint in which hepatocytes are highly proliferative and in the process of remodeling as it transitions from a fetal to mature state. However, in acSRSF1 HKO the knockout out occurs in adults where the hepatocytes are fully mature and quiescent.

Like the SRSF1 HKO model, the acSRSF1 HKO mice also exhibit repopulation of the tissue parenchyma with SRSF1-expressing cells seen at the 4-week timepoint **(Figure 4.2 A)**. This is expected, as the liver is known to be highly regenerative and can replace the dying hepatocytes. However, unlike the SRSF1 HKO mice which spontaneously generate hepatocytes deficient of the Cre transgene, the acSRFS1 HKO mice repopulate with SRSF1-expressing hepatocytes most likely through transdifferentiation of BECs where SRSF1 is intact. This can be seen from the fluorescent imaging which shows SRSF1-expressing hepatocytes originating from the portal triad region **(Figure 4.4 A)**.

It is evident that ablation of SRSF1 hepatocytes is resulting in eventual death of the cell. This can be seen from the severely elevated serum ALT and AST levels in acSRSF1 HKO which is a marker of liver damage **(Figure 4.8)**. Therefore, it should be expected that widespread loss of hepatocyte activity will result in hepatic failure. Indeed, the phenotypic characterization of this model supports this reasoning. Fasting serum levels of glucose and ketone bodies at 4-weeks show striking evidence of hepatic failure. The acSRSF1 HKO mice have decreased fasting blood glucose levels due to the failure

in gluconeogenesis that is required to maintain fasting levels. Furthermore, under fasting conditions when glycogen stores are depleted, the liver tissue begins generating ketone bodies as an alternative fuel source from the metabolism of fatty acids. The liver is the only tissue that can generate ketone bodies. In the acSRSF1 HKO mice, ketone bodies are also severely diminished during the fasting state which further supports that loss of SRSF1 results in hepatocyte death and liver failure.

The liver also plays a major role in lipid and steroid metabolism. Under normal physiological conditions, the liver is involved in packaging lipids into lipoprotein particles for systemic delivery as well as synthesis of cholesterol for the body. It is evident that loss of SRSF1 results in severe accumulation of lipids within the tissue. Furthermore, there is a striking decrease in adipose tissue stores in acSRSF1 HKO mice **(Figure 4.7)**. This is most likely due to deficiency of hepatic function and failure to generate lipoprotein particles for systemic delivery of fatty acids. This signals the peripheral adipose stores to mobilize fatty acids into circulation for uptake and packaging by the liver. However, due to the failure of hepatocytes in acSRSF1 HKO, this results in a futile cycle in which fatty acids are being sent to the liver where it accumulates. This is supported by the normalized weights of the liver and adipose tissue seen in this model which shows increased liver weight ratios with subsequent decrease in adipose weight ratios **(Figure 4.7)**. The liver also plays a major role in synthesizing cholesterol for the body. Additionally, hepatic and serum cholesterol levels are diminished in the acSRSF1 HKO. In conclusion, acute loss of SRSF1 results phenocopies the pathologies observed in SRSF1 HKO and that loss of SRSF1 results in death of hepatocytes leading to liver failure.

4.5 Chapter Figures

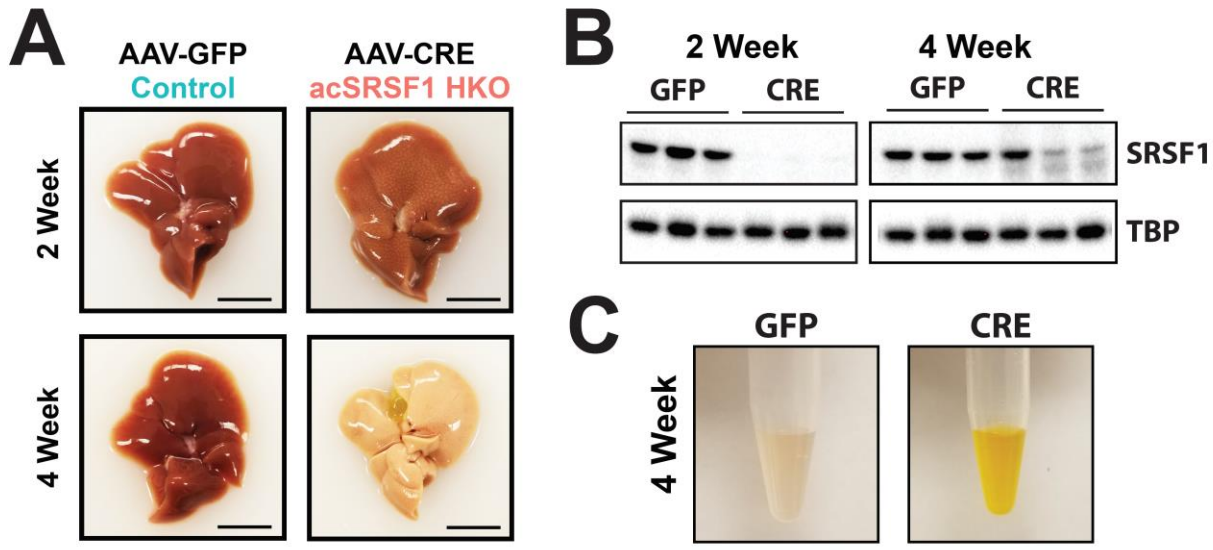


Figure 4.2: ***Acute Knockout of SRSF1 Recapitulates SRSF1 HKO Phenotype***

(A) Gross images of livers from SRSF1 flox/flox mice 2- and 4-weeks post injection with AAV8-TBG-GFP (Controls) or AAV8-TBG-iCre (acSRSF1 HKO). At 2 weeks, acSRSF1 HKO livers are starting to become steatotic. However, at 4 weeks, acSRSF1 HKO livers exhibit severe steatosis with yellow appearance. (B) Western blot showing SRSF1 levels in acSRSF1 HKO. As expected, 2-week mice show complete loss of SRSF1. However, by 4 weeks, acSRSF1 HKO hepatocytes show significant expression of SRSF1 signifying repopulation of the tissue with SRSF1 expressing hepatocytes. (C) Pictures of serum isolated from control and acSRSF1 HKO mice at 4 weeks. The acSRSF1 HKO mice have golden colored serum most likely due to elevated bilirubin levels.

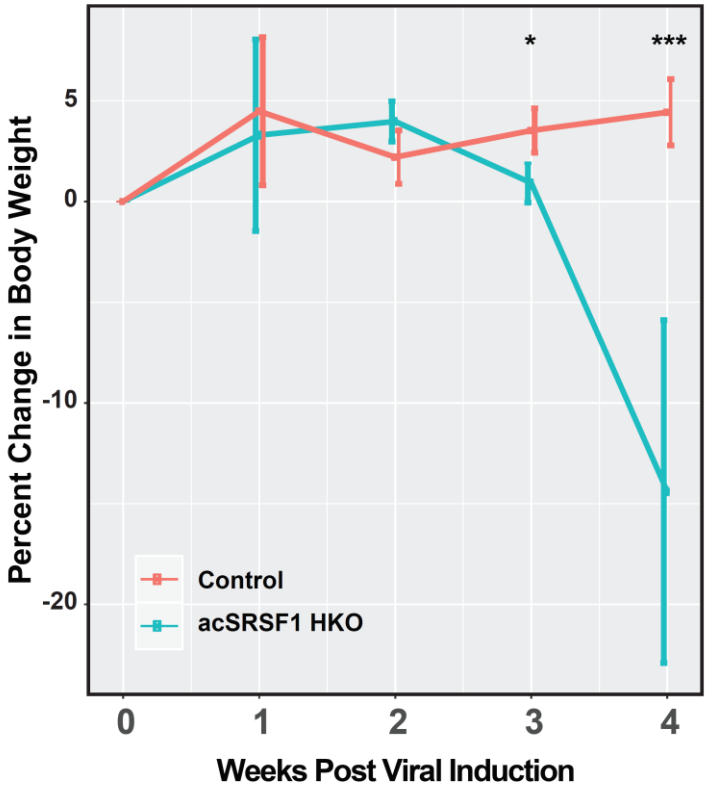


Figure 4.3: ***Acute SRSF1 HKO Mice Exhibit Drastic Loss of Total Body Weight***

Plot showing weekly changes in body weights with respect to initial weight of Control and acSRSF1 HKO post viral induction. AcSRSF1 HKO exhibit a drop in body weight at 3 weeks followed by drastic decrease in body weight at 4 weeks. Acute SRSF1 HKO mice have an average drop of 15% by 4 weeks post viral induction.

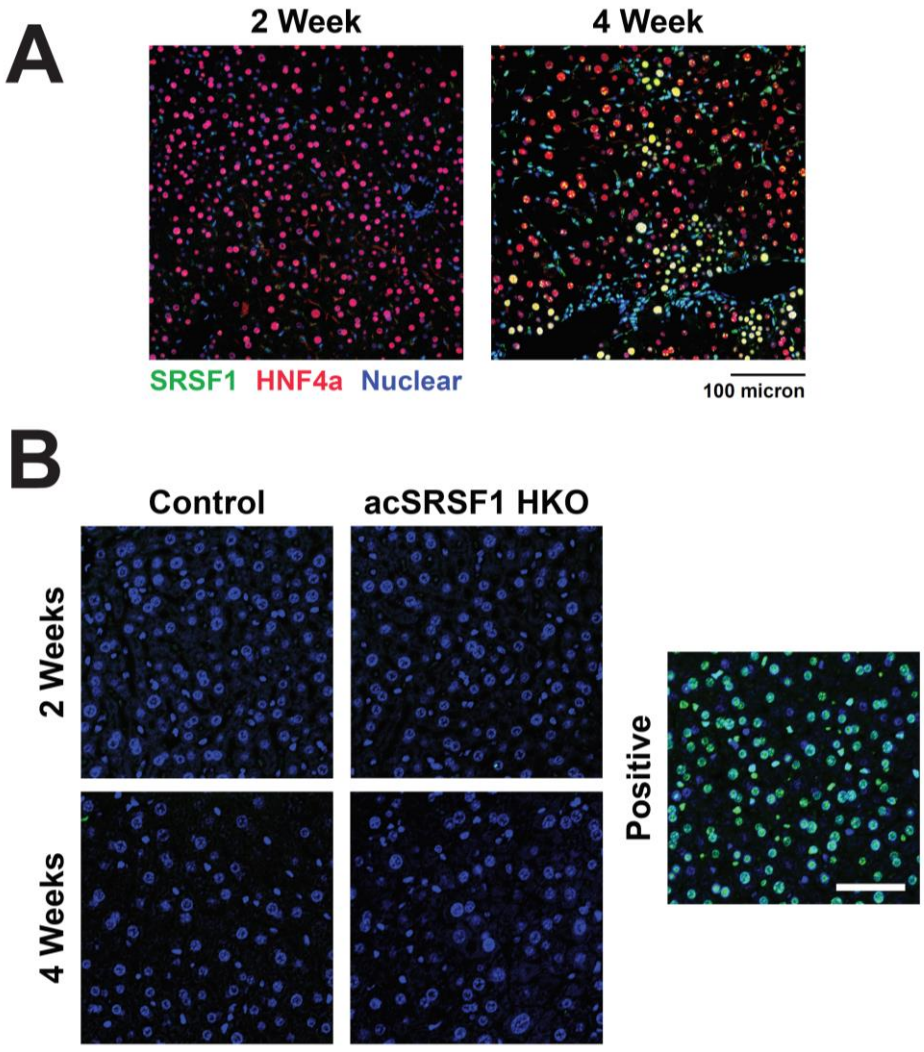


Figure 4.4: ***Eventual Repopulation of SRSF1 Expressing Hepatocytes in acSRSF1 HKO***

(A) Immunofluorescence staining for SRFS1 in acSRSF1 HKO mice at 2- and 4-weeks post viral induction. At 2 weeks, effective knockout of SRSF1 is achieved shown by no SRSF1 signal in hepatocyte nuclei. By 4 weeks, SRSF1 expressing hepatocytes are beginning to repopulate the tissue as seen by the yellow colored nuclei. These SRSF1 expressing hepatocytes originate from the portal triad region. This result agrees with the western analysis showing re expression of SRSF1 at the 4-week timepoint. (B) TUNEL staining of acSRSF1 HKO at 2- and 4-weeks. Staining are negative signifying that these cells do no undergo apoptosis. Scale bar = 100 micron.

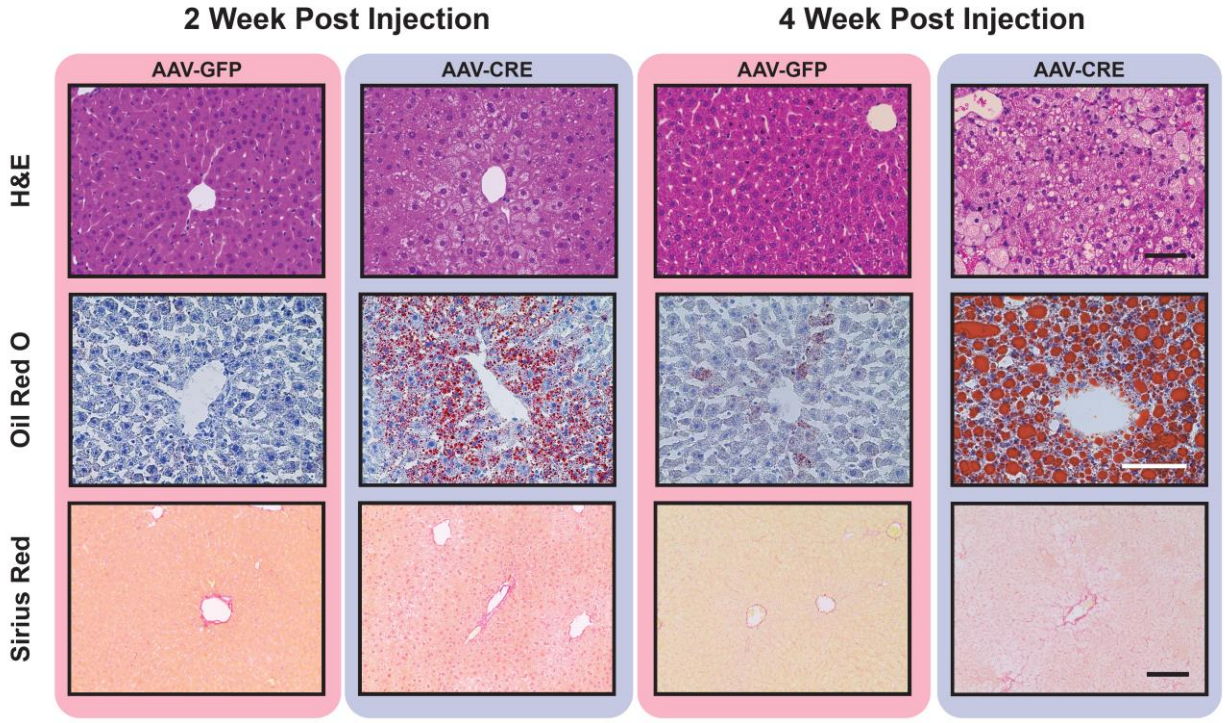


Figure 4.5: ***Histological Analysis of acSRSF1 HKO at 2- and 4-week Timepoints***

To gain insight into the morphological changes within the tissue of acSRSF1 HKO mice, H&E, Oil Red O, and Sirius Red staining were performed. Representative image of each staining for each group at each timepoint are shown (n = 4-5 per group). H&E staining at 2 weeks shows overall tissue architecture is intact with evidence of steatosis. By 4 weeks, the tissue architecture is drastically altered with ballooning cell death. Oil Red O staining shows progressive accumulation of lipids within acSRSF1 HKO livers which agrees with the gross liver images. Sirius Red staining shows no sign of fibrosis at either timepoints.

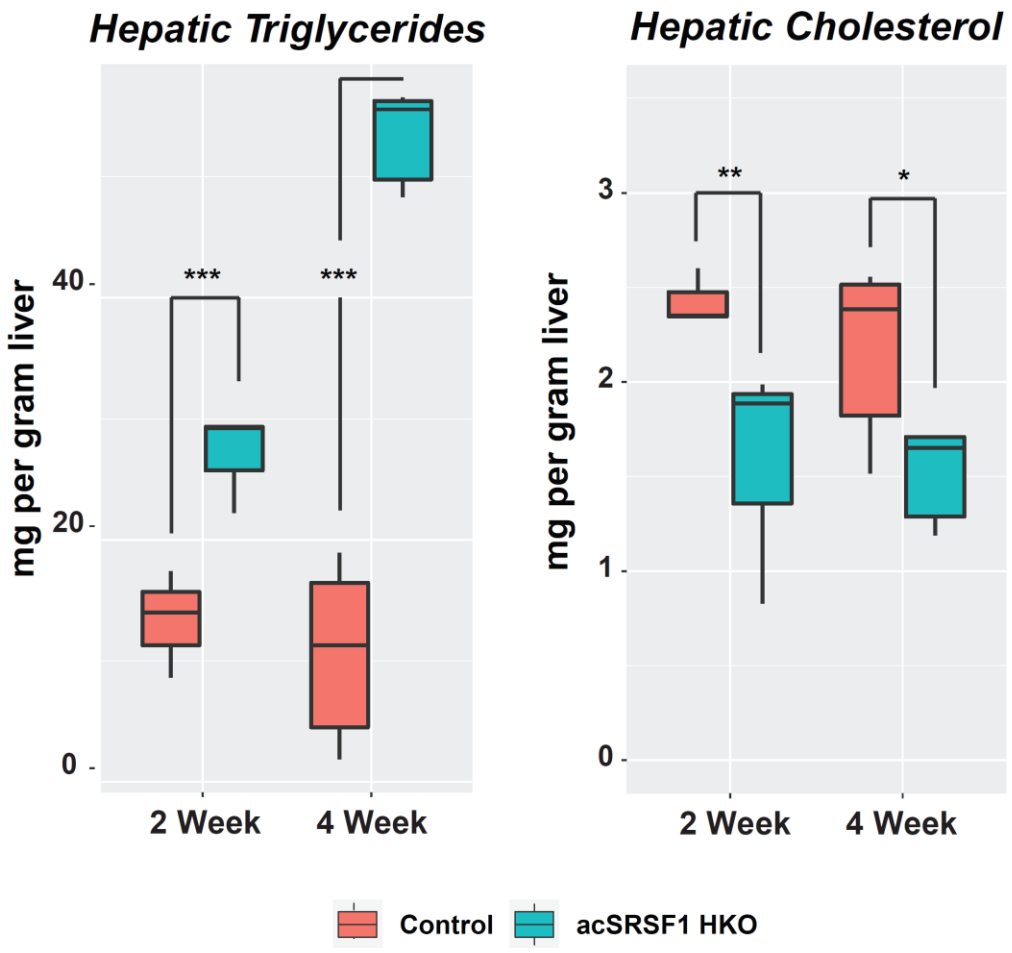


Figure 4.6: ***Hepatic Triglycerides and Cholesterol Levels in acSRSF1 HKO Mice***

Box plots of hepatic triglyceride and cholesterol levels at 2- and 4-week post viral induction timepoints in acSRSF1 HKO mice (n = 6 – 8 mice per group). AcSRSF1 HKO exhibit progressively increasing levels of hepatic triglycerides while cholesterol levels are decreased at both timepoints. Two-way ANOVA statistical test was used to determine significance between 2 groups at 2 timepoints. *P < 0.05, **P < 0.01, ***P < 0.001.

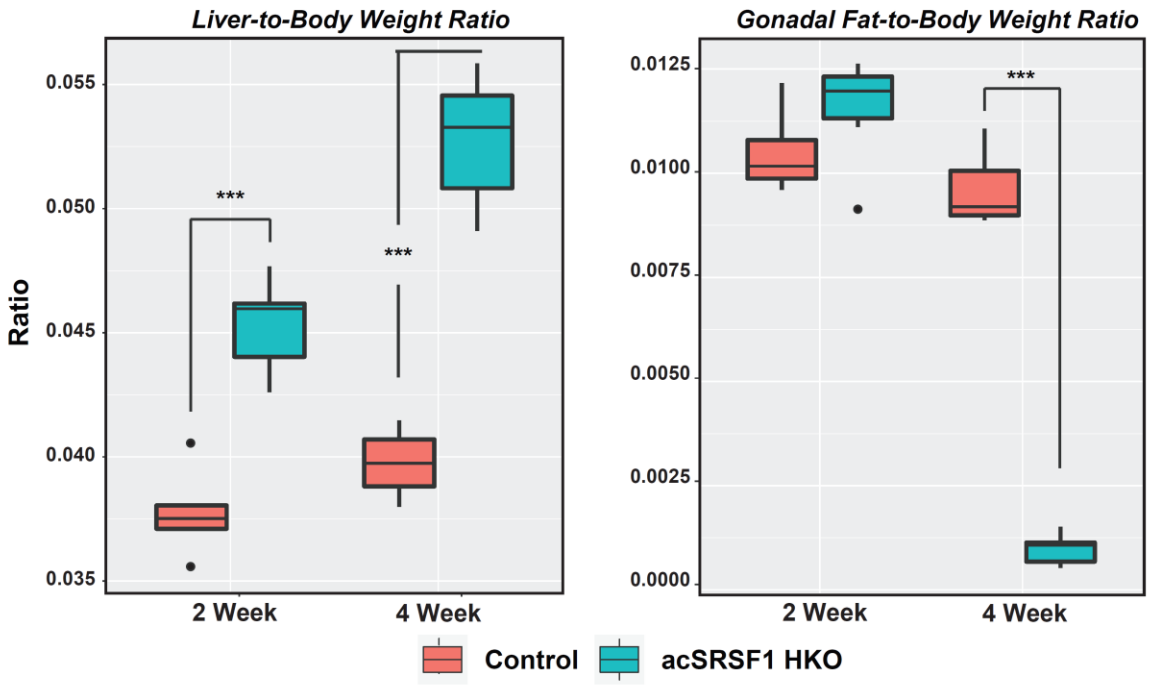


Figure 4.7: ***Liver- and Fat-to-Body Weight Ratio in AcSRSF1 HKO Mice***

Liver weight and gonadal fat weight normalized to total body weight at 2- and 4-week post viral induction timepoints in acSRSF1 HKO mice (n = 6 – 7 mice per group). AcSRSF1 HKO exhibit elevated liver-to-body weight ratios at both timepoints. However, gonadal fat-to-body weight ratios are significantly decreased at 4 weeks in acSRSF1 HKO mice. Two-way ANOVA statistical test was used to determine significance between 2 groups at 2 timepoints. *P < 0.05, **P < 0.01, ***P < 0.001.

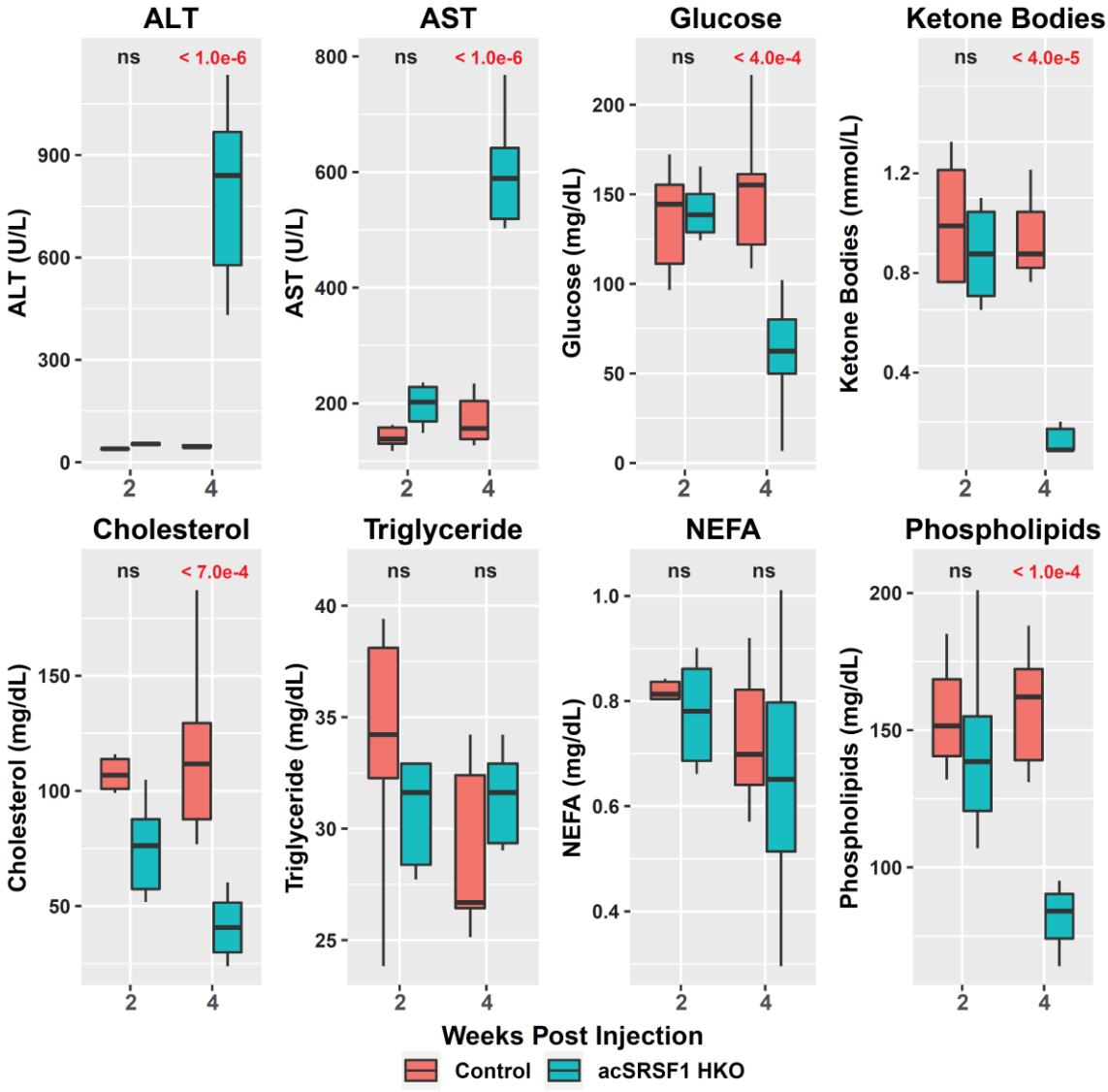


Figure 4.8: ***Serum Profiling in acSRSF1 HKO Mice Reveal Striking Liver Damage***

Serum profiling of multiple serum parameters in 6 hours fasted acSRSF1 HKO mice at both 2- and 4-week timepoints. No significant change is seen in any serum parameter at 2-weeks. By 4 weeks, acSRSF1 HKO exhibit significant liver damage as seen by elevated AST and ALT levels. Furthermore, fasting ketone bodies and cholesterol levels are significantly decreased showing diminished hepatic functionality. Two-way ANOVA statistical test was used to determine significance between 2 groups at 2 timepoints. *P < 0.05, **P < 0.01, ***P < 0.001.

Chapter 5: Loss of SRSF1 Results in R-Loop Accumulation and Widespread DNA Damage

5.1 Introduction

Acute hepatic knockout of SRSF1 results in robust accumulation of lipids within the cells. The acute model recapitulates the pathology observed with the transgenic model, SRSF1 HKO. While severe steatosis is observed by post-natal day 10 in SRSF1 HKO, it takes four weeks to develop steatosis to similar levels in the acSRSF1 HKO. This delayed response in acSRSF1 HKO is most likely due to the difference in developmental timepoint at which the knockout occurs. At early stages, the liver is highly dynamic as it is transitioning from a fetal to a mature state. However, in the acSRSF1 HKO model, knockout is induced in adult livers which consist of fully matured and quiescent hepatocytes. This delayed pathological development provides opportunity to capture SRSF1 deficient hepatocytes before secondary changes begin to take place.

To determine early changes occurring within this model, RNA-Seq was performed two weeks post viral mediated knockout. At this timepoint, robust knockout of SRSF1 has been achieved, however, severe pathological changes have not yet developed. In addition to RNA-seq analysis, an eCLIP-seq for SRFS1 was performed in adult wildtype mice. This technique allows for the determination of direct mRNA targets of RNA binding proteins at the single nucleotide resolution. The goal of performing eCLIP-seq was to further aid in understanding the primary mechanism resulting in development of the observed pathology Specifically, the transcriptome changes determined from RNA-seq can be cross referenced with the eCLIP-seq binding data to identify key players.

5.2 Material and Methods

The following sections provides details of the materials and methodology utilized for the studies performed in this chapter. For details on additional methodology refer to *Section 9.1: Supplementary Protocols.*

5.2.1 RNA-Seq Analysis of acSRSF1 HKO Mice Model

RNA-Seq analysis was performed on isolated hepatocytes from 2-week control and acSRSF1 HKO mice (n = 3 per group) as described in *Section 3.2.1*. The data was analyzed using workflow described in *Section 3.2.2*.

5.2.2 eCLIP-seq Library Preparation, Sequencing, and Data Processing

Hepatocytes were isolated from livers using the Two-Step Collagenase Perfusion technique as described in *Section 9.1.2*. Isolated hepatocytes suspended in PBS were then exposed to 400 mJ/cm^2 of 254 nm UV radiation to crosslink and stabilize RNA binding protein (RBP)–RNA interactions. Subsequent immunoprecipitation of SRSF1-RNA complexes, RNA isolation, library preparation and sequencing were performed as previously described (Van Nostrand et al. 2016). Raw reads were mapped to the mm9 mouse genome using STAR and peak calling was performed using the publicly available tool CLIPper (https://github.com/YeoLab/clipper). This tool provided identified peaks in a BED12 format. This data was further processed and analyzed using custom python and R scripts.

5.2.3 Dot Blot Assays for DNA-RNA Hybrid and PolyA mRNA

DNA-RNA hybrid dot blot assay was adapted from a previously published report (Morales et al. 2016). Briefly, DNA was isolated from approximately 50 mg of liver tissue or snap frozen cells using DNeasy Blood & Tissue Kit (Qiagen) using the manufacturer's

protocol. For each sample 250 ng of DNA was digested with 5 Units of RNase H (NEB) as control. The slot blot apparatus was setup by placing 2 layers of filter paper soaked in PBS with a Hybond^{TM}-N+ membrane (Amersham) layered on top before clamping down the well block. 250 ng of DNA was diluted with PBS to 200 µL and then transferred to the apparatus well. Vacuum suction was applied to the apparatus outlet to allow for suctioning of well volume through the membrane layers. The membrane was then auto cross-linked with UV (254 nm, 1200 mJ/cm^2). The membrane was then air-dried before beginning blocking with 5% nonfat dry milk and 0.1% Tween 20 (TBST). Membrane was then incubated overnight at 4°C with S9.6 antibody diluted in blocking buffer. Finally, it was probed with Mouse-HRP secondary antibody and developed as described in *Section 9.1.3.* Refer to *Section 9.1.9* for details on antibody source and dilutions.

Dot blot analysis for polyA mRNA was performed on total RNA isolated from approximately 50 mg of snap frozen liver tissue using the RNeasy kit following standard protocols. 250 ng of total RNA was blotted onto membrane as described previously for DNA-RNA hybrid assay. Membrane was UV crosslinked and air dried before continuing. The blot was prehybridized with ULTRAhyb-Oligo Hybridization Buffer (Thermo Fisher Scientific) for 10 minutes at 42 °C with gentle agitation in a hybridization oven. PolyA mRNA was detected using biotinylated oligo-dT probes (Promega #Z5261). Probe was diluted in ULTRAhyb-Oligo Hybridization Buffer (0.5 µL of 5 pmol/uL per mL of buffer). Blot was incubated with probe solution at 42 °C for 1 hour. Blot was then washed twice with ULTRAhyb-Oligo Hybridization Buffer. Chemiluminescent detection was performed using the Chemiluminescent Nucleic Acid Detection Module Kit (Thermo Fisher Scientific # 89880) following manufacturers protocol.

5.3 Results

5.3.1 RNA-seq Analysis of acSRSF1 HKO Reveal Transcriptome-wide Defects

Like SRSF1 HKO, RNA-seq was performed in acSRSF1 HKO model at the 2-week timepoint. This timepoint was chosen as the severe pathological changes have not yet developed. Differential gene expression and splicing analysis reveal a substantial number of changes to the transcriptome. With regards to expression, over 3,700 genes were found to be differentially expressed **(Figure 5.1 A)**. This corresponds to nearly 30% of the hepatocyte transcriptome. Furthermore, a greater proportion of genes are upregulated with 2,525 genes which accounts for nearly two-thirds of all DEGs.

In addition to gene expression changes, acSRSF1 HKO showed significant changes in differential splicing **(Figure 5.1 B)**. A total of 2,996 splicing events were found to be significantly changing in PSI between control and acSRSF1 HKO. Like SRSF1 HKO, the event type with the largest representation was SE with a total of 2040 events. Furthermore, a larger proportion of events exhibited decreases inclusion with loss of SRSF1. For instance, about 70% of the differentially spliced SE events were found to have decreased inclusion.

It is known that differential splicing can affect the stability of transcripts which can shift steady state abundances of transcripts. To determine if splicing and gene expression are dependent, an overlap between the two sets was performed **(Figure 5.2)**. A total of 248 genes were found to be differentially spliced and expressed in acSRSF1 HKO. The calculated p-value for the overlap was 0.30 and therefore was not found to be significant. Furthermore, of the overlapped genes, about 54% of the genes exhibited downregulation. Therefore, it appears the differential splicing and gene expression are independent.

5.3.2 Splicing Defects in acSRSF1 HKO are Indirect Effects of SRSF1 Loss

Differential splicing analysis in the acSRSF1 HKO revealed widespread defects in exon usage. However, to determine which of the differentially spliced exons are directly regulated by SRSF1, an eCLIP-seq experiment was performed in hepatocytes isolated from wildtype mice. This analysis revealed a total of 5,727 significant unique peaks for SRSF1 in hepatocytes **(Figure 5.3 A)**. Most binding of SRSF1 is found within exons, followed by introns and then exon-intron boundary. To determine the type of genes on which SRSF1 binds, a gene ontology was performed on all the genes with SRSF1 binding **(Figure 5.3 B)**. The analysis shows enrichment for genes involved in RNA binding and metabolic processes such as lipoprotein particle receptor binding and sterol transport.

To further investigate the direct effects of SRSF1 ablation on the transcriptome, the identified peaks was overlapped with the genes changing in expression or splicing **(Figure 5.4)**. Surprisingly, of all the identified SRSF1 binding peaks, about 7% were found to have binding on exons with differential splicing in the acSRSF1 HKO. This is lesser than the overlap with genes found to be downregulated in expression with 17.4% of binding peaks. This suggests that the widespread splicing defects observed in acSRSF1 HKO is independent of direct SRSF1 binding. It has been seen that RNA binding proteins can have preferential binding to certain regions of mRNA, such as the CDS, 5’ UTR, and 3’ UTR. Binding in particular regions can confer different regulatory effects such as stability or transportation. Determination of SRSF1 binding propensity regarding region type shows that SRSF1 binds primarily to CDS. This suggests that SRSF1 does not have varying propensity by type since CDS make up a larger portion of mRNA. Furthermore,

the proportion of SRSF1 binding by mRNA region does not change drastically for genes that are changing in expression or splicing **(Figure 5.5)**.

Analysis of the spatial distribution of SRSF1 binding was performed on the differentially skipped exon set from acSRSF1 HKO **(Figure 5.6)**. Since it is likely that the differentially skipped exons in the acSRSF1 HKO are regulated by SRSF1, it is expected that these events would exhibit enrichment of SRSF1 binding peaks. Shockingly, compared to a background constitutive exon set, there is a striking de-enrichment of SRSF1 binding in the differentially spliced exons. This provides further support that the splicing defects detected in acSRSF1 HKO is independent of SRSF1 known splicing activity. Motif analysis was also performed on the identified SRFS1 binding peaks **(Figure 5.7 A)**. The peaks enriched for the canonical GAAGAA motif that has been found in previous SRSF1 binding studies. The spatial distribution of this motif was found to be enriched within exonic regions **(Figure 5.7 B)**. A second motif with the sequence CUA/UCA was also found to be enriched in the SRSF1 binding peaks. However, strong spatial enrichment of this motif within exon regions was not found. Instead enrichment was found in intronic regions, however, the functionality of this motif is not clear.

An interesting observation noticed of the genes with SRSF1 binding is that the distribution of expression changes reveals there to be an overall decrease in abundance in acSRSF1 HKO model **(Figure 5.8)**. A possible explanation of this finding is that binding of SRSF1 to these genes confers additional stability. Therefore, loss of SRSF1 leads to destabilization of the mRNA resulting in decreased steady state levels of the transcript. This is possible because it is known that SRSF1 is involved in forming EJC complex which facilitates compaction of mRNA and provides stability.

5.3.3 Gene Ontology Analysis Reveals DNA Damage Response in acSRSF1 HKO

To investigate the gene regulatory networks affected in acSRSF1 HKO, gene ontology analysis was performed on the differential gene expression and splicing data. The resulting gene ontology network revealed enrichment of multiple biological processes **(Figure 5.9)**. Like SRSF1 HKO, acSRSF1 HKO displayed a strong upregulated enrichment for processes involved in immune response. This suggests that even at this early timepoint, loss of SRSF1 triggers the hepatocytes to activate signals for recruitment of inflammatory cells. It is possible that these signals are further exacerbated by the introduction of viral vectors into the liver. A strong enrichment is also observed for regeneration response. At this timepoint, the hepatocyte functions are beginning to decrease which is likely triggering an early compensatory response to begin repopulating the tissue with additional hepatocytes. Interestingly, the network map also shows a strong upregulation for processes involved in DNA repair and cell cycle in the setting of acSRSF1 HKO. These terms also have a strong overlap with genes that were also found to have differentially spliced exons. It is known that SRSF1 is plays a role in maintaining genome stability. Therefore, loss of SRSF1 is resulting in DNA damage which is triggering a DNA damage repair response.

5.3.4 AcSRSF1 HKO do not Exhibit Activation of the Unfolded Protein Response

A possible mechanism of the pathological findings observed in the SRSF1 HKO models is that loss of SRSF1 is resulting in accumulation of unfolded proteins. Specifically, absence of SRSF1 is leading to transcription of transcripts deficient of constitutive exons which will eventually be translated into aberrant proteins. These aberrant proteins will then lead to accumulation of improper folding and activation of the

unfolded protein response. To investigate this hypothesis, heatmaps for UPR responsive genes were generated **(Figure 5.10 A)**. It is evident from the heatmaps that there is a strong downregulation of UPR genes. Furthermore, a canonical response to the UPR is the splicing of the Xbp1 gene. Presence of unfolded proteins within the ER lumen results dimerization of IRE1 and subsequent unconventional splicing of Xbp1 leading to the formation of a short isoform (Dara, Ji, and Kaplowitz 2011). The splicing of this exon is unchanged in acSRSF1 HKO mice **(Figure 5.10 B)**. Finally, UPR also leads to increased levels of CHOP, a UPR downstream effector protein and transcription factor (Samali et al. 2010; Oslowski and Urano 2011). The levels of this protein are also unchanged in acSRSF1 HKO hepatocytes **(Figure 5.10 C)**. Taken together, it is highly unlikely that loss of SRSF1 is resulting in accumulation of unfolded proteins leading to proteotoxic affects.

5.3.5 DNA Damage in SRSF1 HKO is Likely Mediated by R-Loop Accumulation

Gene ontology analysis revealed that loss of SRSF1 triggers a DNA damage repair response. It is well established that SRSF1 is involved in maintaining genome integrity. This is due to SRSF1's ability to reduce the formation of R-loops which are DNA-RNA hybrid structures. Loss of SRSF1 is known to cause accumulation of these structures as genes are being transcribed. This overwhelming formation of R-loops leads to DNA double strand breaks as the cell is not able to resolve them. To determine if DNA damage is present in the acSRSF1 HKO mice, immunofluorescent staining for γH2A.X, a well-known marker of DNA damage, was performed on tissue from acSRSF1 HKO (Mah, El-Osta, and Karagiannis 2010). In agreement with this hypothesis, robust signal for γH2A.X is seen in the acSRSF1 HKO mice **(Figure 5.11)**. Furthermore, this signal is also observed in SRSF1 HKO at the 10-day and 1-month timepoint. Next, to confirm that this

DNA damage is likely a result of increased accumulation of R-loops. A S9.6 dot blot assay was performed using DNA isolated from acSRSF1 HKO. The S9.6 antibody specifically binds to DNA-RNA hybrids. The dot blot shows robust signal in SRSF1 knockout samples for DNA-RNA hybrids in both the acSRSF1 HKO and SRSF1 HKO models **(Figure 5.13 A)**. It has been shown that the widespread accumulation of R-loops results in blockage of transcription. It can be expected that blockage of transcription will result in decrease production of nascent mRNA and therefore a decrease in steady state abundance of polyA mRNA.

To determine if inhibition of global transcription may be occurring in SRSF1 knockout hepatocytes, a dot blot using a polydT probe was performed on total RNA to determine steady state levels of polyA mRNA **(Figure 5.13 B)**. Indeed, a significant decrease in total polyA mRNA is observed in SRSF1 HKO samples compared to controls. It should be noted that equivalent amounts of total RNA were blotted in this assay. Therefore, this assay does not provide information regarding absolute quantities of polyA mRNA. The levels of polyA mRNA relative to total RNA is decreased in SRSF1 HKO hepatocytes. Since mRNA makes up about 3% - 5% of total RNA, this assay provides insight into the differences in abundance of polyA mRNA relative to the total RNA pool. While it is expected that relative abundance of polyA mRNA relative to the total RNA pool will decrease in the setting of decreased global transcription, this observation would be expected in multiple cases. For instance, increased degradation of polyA mRNA would also be expected to show similar findings.

5.4 Conclusion

Extensive characterization of SRSF1 HKO mice demonstrated that loss of SRSF1 activity in hepatocytes results in cellular death. This is not unexpected since previous reports have shown that knocking out SRSF1 is embryonically lethal. However, to determine the molecular mechanisms leading the demise of the cell was not clear. Significant effort was placed in understanding the transcriptome changes occurring in the model using techniques such as RNA-seq and eCLIP-seq. However, these efforts did not provide the great insight. Analysis of the RNA-seq and SRSF1 binding data showed that most of the observed transcriptome changes were not a result of SRSF1's splicing deficiency **(Figure 5.5 and 5.7)**. The motivation for utilizing these methodologies was based on the hypothesis that loss of SRSF1 was leading to misregulation of a specific set of transcripts needed for the viability of the cell. In retrospect, this hypothesis was completely incorrect. Fortunately, analysis of the acSRSF1 transcriptome shed light into the possibility of DNA damage response and repair **(Figure 5.10)**.

Further investigating this possibility showed that DNA damage was occurring in the SRSF1 HKO models **(Figure 5.12)**. In fact, multiple reports have shown that loss of SRSF1 activity results in genome instability. It was shown that deficiency of SRSF1 causes a robust accumulation of R-loops which are three-stranded nucleic acid structures composed of DNA-RNA hybrids (X. Li and Manley 2005; X. Li, Niu, and Manley 2007). Formation of these structures occurs at some basal frequency in all cells during transcription, however, a healthy cell can resolve them without leading to DNA damage. But, conditions in which the frequency of R-loop formation increases, the cell can be overwhelmed and fails to resolve these structures leading to DNA double strand breaks.

In addition to all the activities of SRSF1, it is also involved in preventing the formation of R-loops. This is presumably due to the co-transcriptional binding of mRNA preventing it from annealing back to the template DNA (Tous and Aguilera 2007).

Revelation of this understanding completely changed the hypothesis regarding the molecular mechanism resulting in death of SRSF1-deficient hepatocytes. It is well established, that the DNA damage resulting from SRSF1 knockout is most likely due to R-loop formation. To confirm if SRSF1 knockout model also show accumulation of R-loops, a dot blot assay using the S9.6 antibody was used which detects DNA-RNA hybrids. Indeed, this assay showed a striking signal for R-loop presence in DNA isolated from the SRSF1 knockout mice **(Figure 5.13 A)**. Therefore, the demise of the SRSF1-deficient hepatocytes is most likely due to the extensive and unresolvable DNA damage induced by R-loop accumulation. With the widespread accumulation of R-loops in SRSF1 HKO, it is expected that global transcription will be blocked. To shed some light into this possibility, relative polyA mRNA levels were quantified by a dot blot northern assay using a polydT probe **(Figure 5.13 B)**. Results from this assay revealed a significant decrease in relative steady state levels of polyA mRNA in SRSF1 HKO.

It should be noted that this assay measures the steady state levels of polyA mRNA and not nascently transcribed mRNA. The hypothesis is that loss of SRSF1 is resulting in a decrease of global transcription leading to subsequent decrease of total RNA. Since non polyadenylated ribosomal RNA makes up about 80% of total RNA, the decreased signal for polyA mRNA signifies that relative abundance of polyA mRNA to total mRNA is decreasing. If loss of SRSF1 is indeed resulting in decreased transcription globally, the decreased polyA signal seen in this assay suggests that the polyA mRNA exhibit higher

turnover rates in comparison to ribosomal RNA. This higher turnover rate will result in greater difference in relative polyA mRNA abundance to total RNA which would allow for differences in the polyA mRNA signal to be detected. However, this result would also be seen in the case if SRSF1 knockout was inducing higher turnover rates of polyA mRNA. To truly prove that global mRNA transcription is decreased in SRSF1 HKO, another technique must be utilized such as metabolic labeling of nascent mRNA.

Finally, it should be stated that the transcriptome data did provide strong evidence that the pathological effects of SRSF1 deficiency is not a result of toxic protein accumulation. A possible hypothesis of toxic effects of SRSF1 deficiency is due to the generation of mispliced transcripts which translate into aberrant proteins that do not fold properly. This would result in the activation of the Unfolded Protein Response or UPR. Multiple pieces of evidence in the SRSF1 HKO model reject this mechanism. This includes gene expression data which shows downregulation of UPR genes, no induction of CHOP levels by western blot analysis and no change in XBP1 splicing **(Figure 5.11)**. This is understandable since the loss of SRSF1 is leading to damage at the transcription level, before any misplicing and translation can occur for the UPR to be triggered.

5.5 Chapter Figures

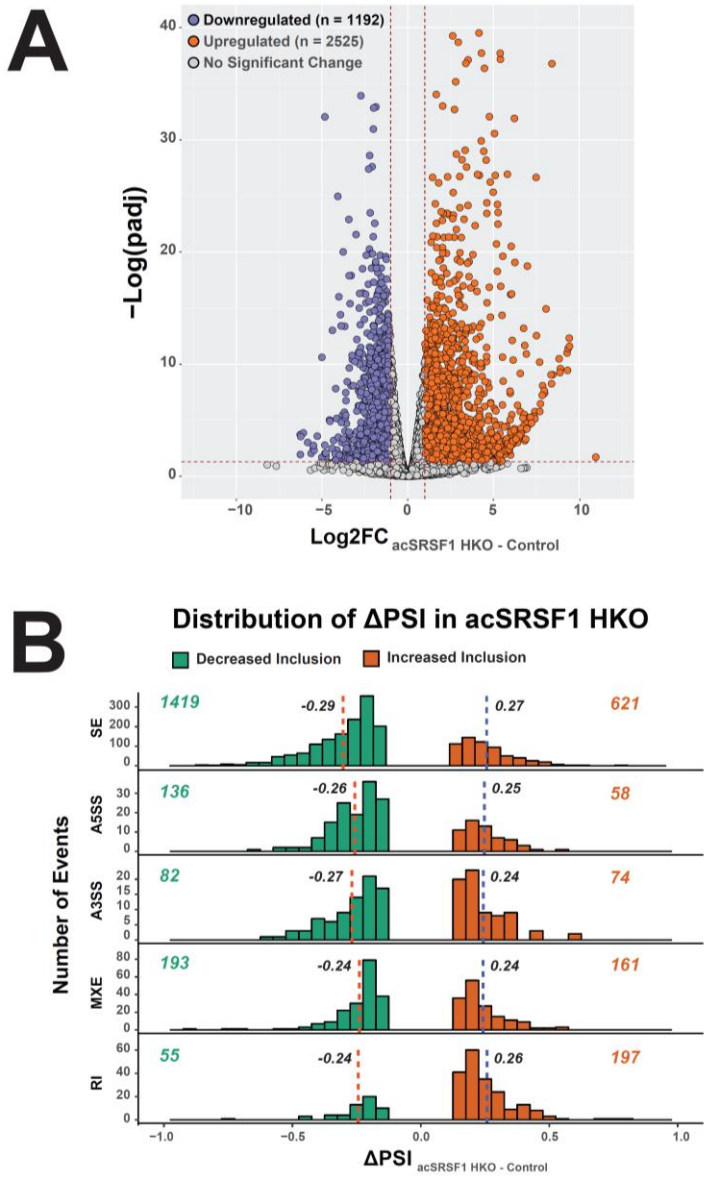


Figure 5.1: ***Gene Expression and Exon Splicing Distributions in acSRSF1 HKO***

(A) Volcano plot showing changes in mRNA abundances from RNA-seq (n = 3 samples per condition) in hepatocytes from acSRSF1 HKO 2 weeks post viral induction. A total of 3717 genes are changing in gene expression with over two-thirds being upregulated. (B) Histogram showing distribution of calculated ΔPSI from RNA-seq for each splicing event type. The dotted line marks the median of the distribution. A larger percentage of exons have decreased inclusion.

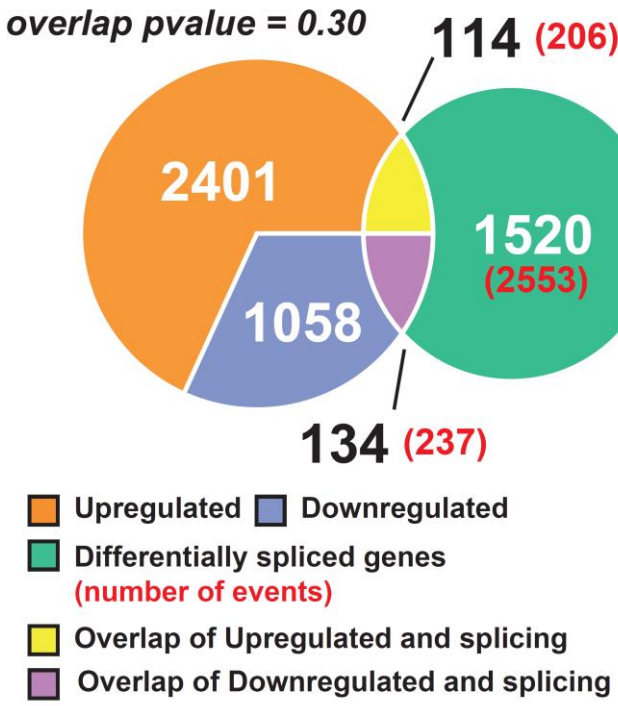
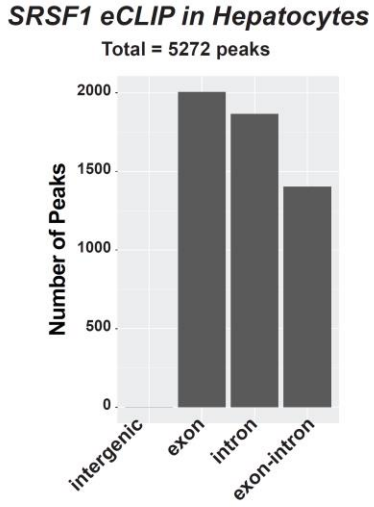


Figure 5.2: ***Overlap Between Gene Expression and Splicing in acSRSF1 HKO***

Venn diagram showing breakdown of the number of genes with differential gene expression and splicing and the overlap between the two sets. The red values signify the number of events. The overlap between the two sets are not significant (p value > 0.05), thus suggesting the changes two forms of regulation are independent.

A



B

Term	Rank	P-value
RNA binding (GO:0003723)	1	2.76E-23
lipoprotein particle receptor binding (GO:0070325)	2	1.59E-06
phosphatidylcholine-sterol O-acyltransferase activator activity (GO:0060228)	3	1.62E-06
cadherin binding (GO:0045296)	4	3.10E-05
sterol transporter activity (GO:0015248)	5	4.20E-05
adenyl ribonucleotide binding (GO:0032559)	6	4.86E-05
lipase inhibitor activity (GO:0055102)	7	5.76E-05
ADP binding (GO:0043531)	8	8.20E-05
cholesterol transporter activity (GO:0017127)	9	1.06E-04
mRNA binding (GO:0003729)	10	1.92E-04

Figure 5.3: ***SRSF1 eCLIP Peak Summary and Gene Ontology Analysis***

SRSF1 eCLIP-seq was performed using hepatocytes isolated from C57BL/6j mice. Peaks were then identified using Clipper and then annotated using custom python scripts. (A) A graph showing the number of SRSF1 binding peaks found at the indicated pre-mRNA region. (B) Gene Ontology analysis of all genes with SRSF1 binding peaks. Genes with SRSF1 binding showed enrichment for RNA binding and lipoprotein particle.

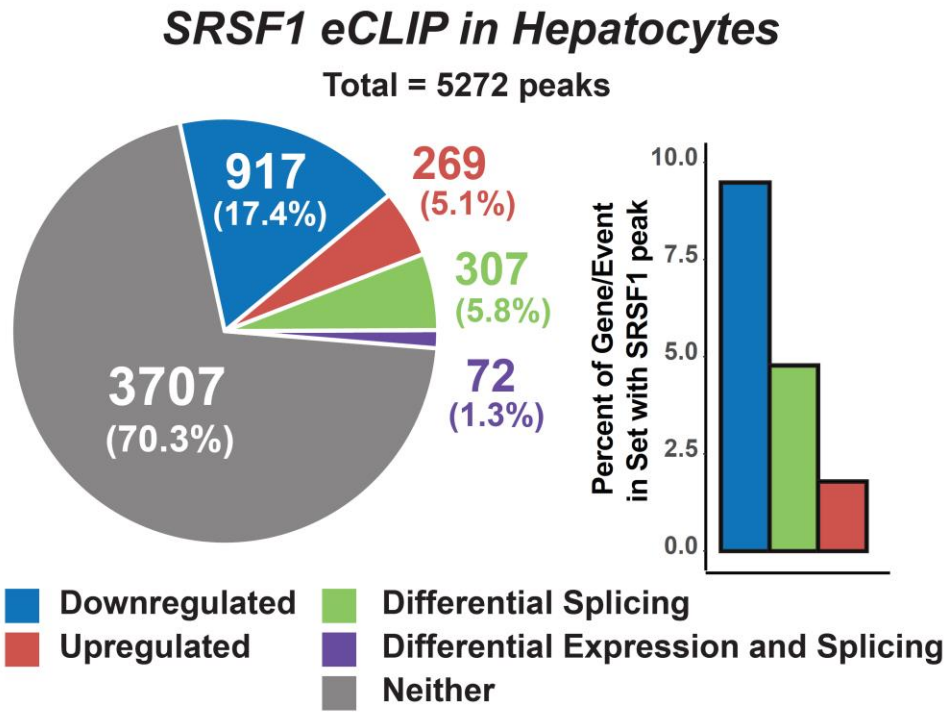


Figure 5.4: ***Overlap of SRSF1 Binding Peaks with SRSF1 Regulated Genes***

Venn diagram showing the breakdown of SRSF1 binding peaks overlapping with genes changing in expression or splicing in the acSRSF1 HKO model. A large percentage (~70%) of SRSF1 binding peaks are localized on genes which were not found to be changing in expression or splicing in the acSRSF1 HKO. The second largest fraction (~17%) of SRSF1 binding peaks were associated with genes found to be downregulated. This is followed by peaks associated with differential splicing and then upregulated genes (5.8% and 5.1%, respectively). This pattern also holds for the inverse relationship. Of the SRSF1 regulated gene sets, genes downregulated in acSRSF1 HKO had the highest percentage of genes with SRSF1 binding. This was followed by genes with differential splicing and then upregulated genes.

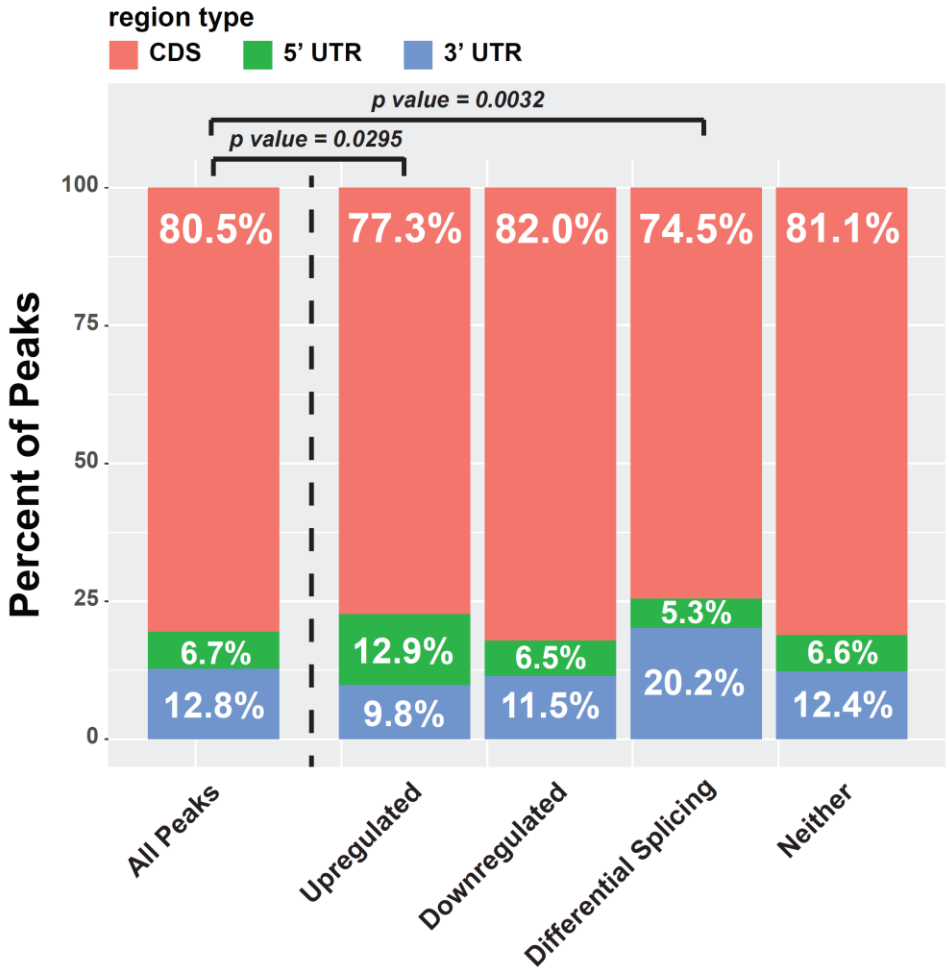


Figure 5.5: ***Breakdown of SRSF1 Binding Peaks by Localization on mRNA Regions***

The three main regions of a transcript include 5' UTR, CDS, and 3' UTR. Of the three regions, the CDS often takes the largest portion of a transcript. To understand if SRSF1 preferentially bound one region over the other, a breakdown of the SRSF1 binding peaks by the associated mRNA region was determined. This analysis was then further performed on the sets of peaks bound to genes that were regulated in acSRSF1 HKO. As expected, for all peaks SRSF1 was found to bind to primarily the CDS region. The ratio of binding peaks by region do not drastically shift based on association with regulated genes. Although peaks on genes with differential splicing did show significantly higher portion localized on the 3' UTR.

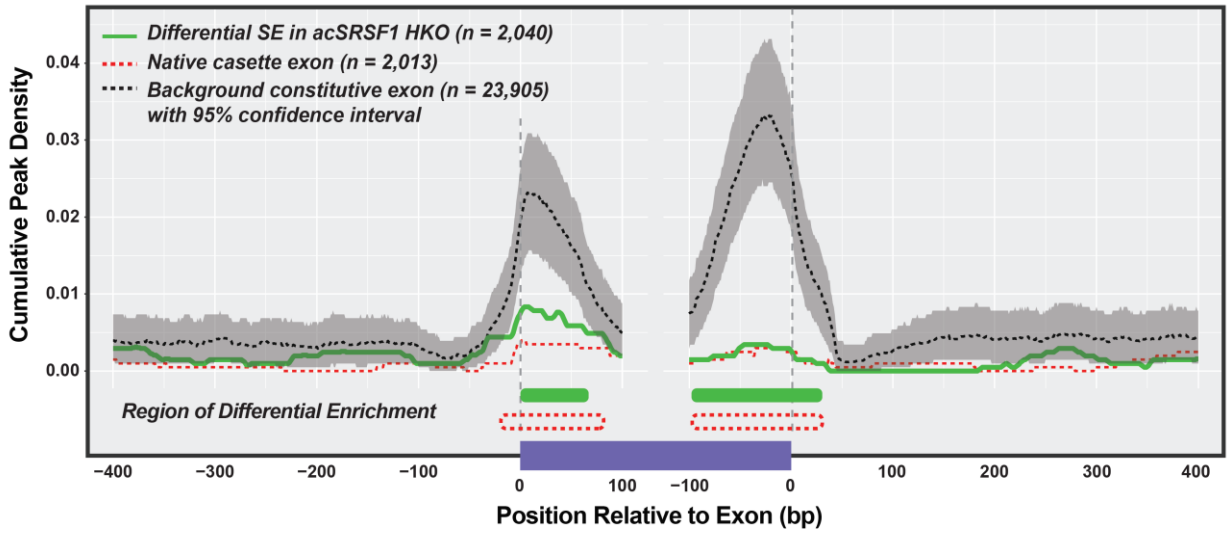


Figure 5.6: ***Spatial distribution of SRSF1 Binding Peaks on Differentially Spliced Cassette Exons in acSRSF1 HKO***

Cumulative peak density plot of SRSF1 binding peaks on the 2,040 differentially spliced cassette exons (green line) in acSRSF1 HKO model. A set of 23,905 constitutive exons were used as background. The background set was sampled 1,000 to determine the 95% confidence interval of peak density at each nucleotide position. The background set shows strong enrichment at exon boundaries. It is apparent that the differentially spliced exons show a strong decreased peak density compared to the background. Furthermore, a native cassette exon set was also used for comparison which consists of exons found to be alternative (0.05 < ΔPSI < 0.95) but did not change significantly upon SRSF1 knockout. These native exons set also showed decreased density of SRSF1 peaks.

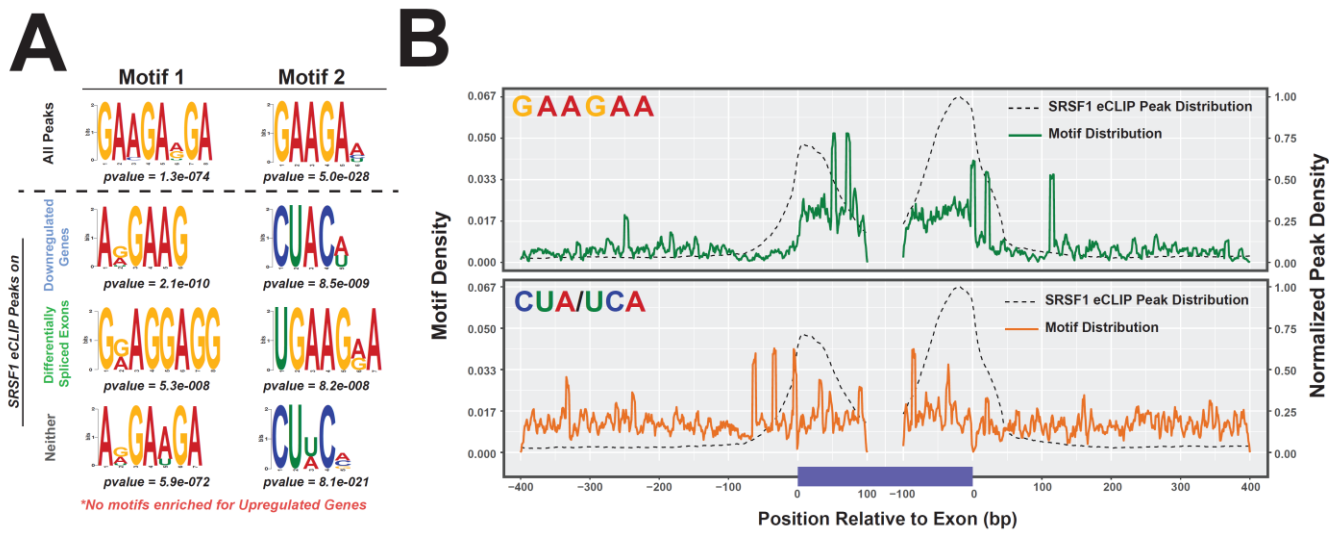


Figure 5.7: ***Motif Analysis of SRSF1 Binding Peaks in Hepatocyte eCLIP-seq Dataset***

(A) Motif enrichment analysis was performed on SRSF1 binding peak regions using MEME suite. Two highly enriched motifs are presented for each set type. For all SRSF1 peaks the canonical GAAGAA motif was enriched. This was the case for binding peaks associated to regulated gene sets. The second motif enriched in the downregulated gene set included a CUAUCA motif. (B) Cumulative spatial distribution of the enriched motifs relative to exon coordinates. Enrichment of the GAAGAA motif was found within exon region. The CUA/UCA motif showed density spikes within the intronic region at the 5' exon boundary.

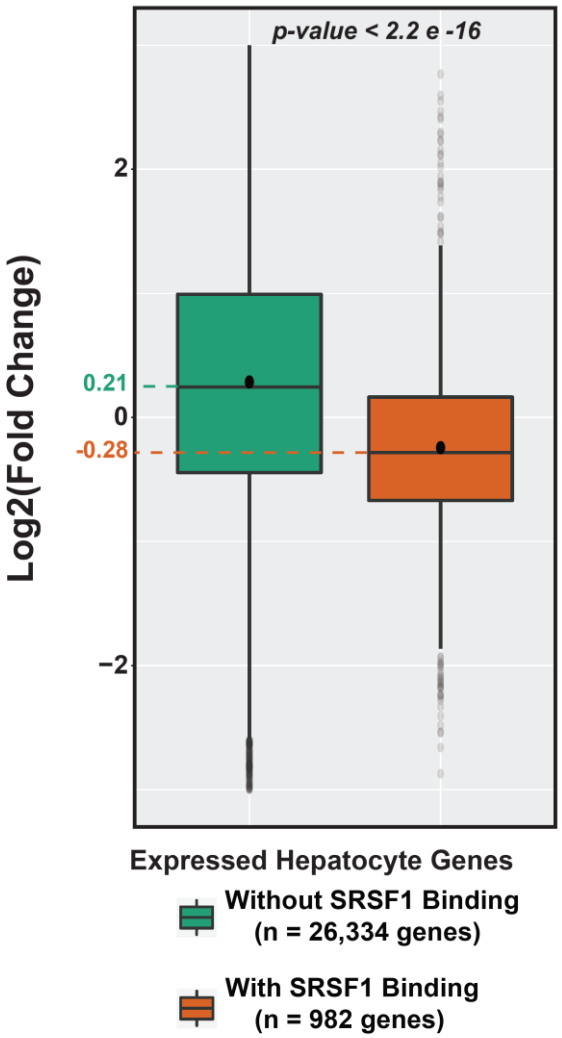


Figure 5.8: ***Genes with SRSF1 Binding Exhibit Decreased Abundance in acSRSF1 HKO***

Box plots showing distribution of Log2 fold change values for genes with and without SRSF1 binding. A total of 982 genes showed SRSF1 binding from eCLIP-seq. Genes with SRSF1 binding exhibited decreased transcript abundances in acSRSF1 HKO in comparison to genes which showed no binding. A possible explanation for this observation is that SRSF1 provides additional stability to transcripts. However, with absence of SRSF1 in the knockout model, these set of mRNA are destabilized and have decreased steady state levels..

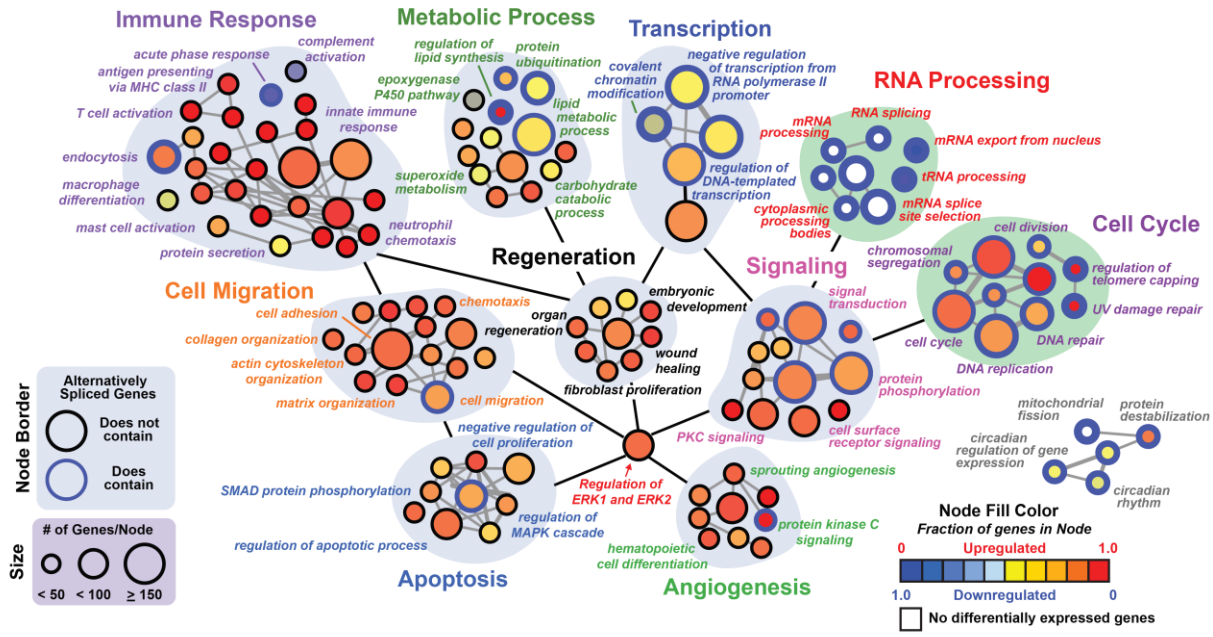


Figure 5.9: ***Gene Ontology Network Map of Genes with Differential Expression and Splicing in acSRSF1 HKO Model***

A gene ontology network map representing the gene ontology terms enriched in the combined set of upregulated, downregulated, and differentially spliced genes. The size of the node provided the number of genes associated with the node. The node color represents the fraction of genes in the node either upregulated or downregulated. Finally, nodes with blue outline represent nodes that contain genes that are differentially spliced. A significant fraction of nodes show that upregulated genes enrich for terms involved in immune response and regeneration. Furthermore, an enrichment for terms involved in DNA damage response and cell cycle shows overlap between increased expression and splicing.

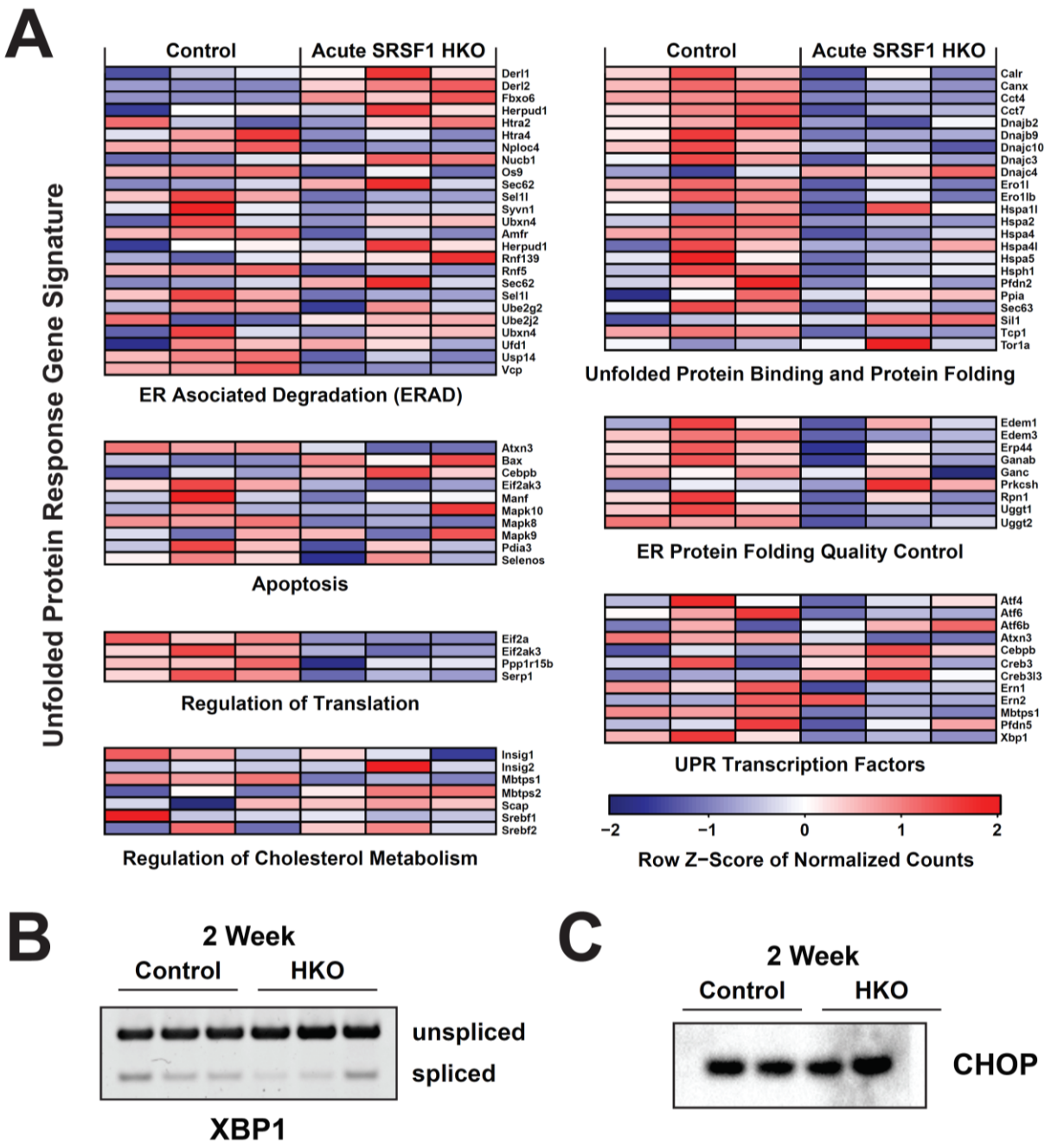


Figure 5.10: ***Acute SRSF1 HKO do not Exhibit an Unfolded Protein Response***

(A) Heatmap showing expression of genes involved in the UPR response in acSRSF1 HKO model. Many genes are downregulated with no evidence of induction of expression. (B) Splice gel of the canonical XBP1 splicing event responsive to UPR. No change in splicing is observed in acSRSF1 HKO. (C) Western blot analysis for CHOP expression in acSRSF1 HKO. No significant change in expression is observed.

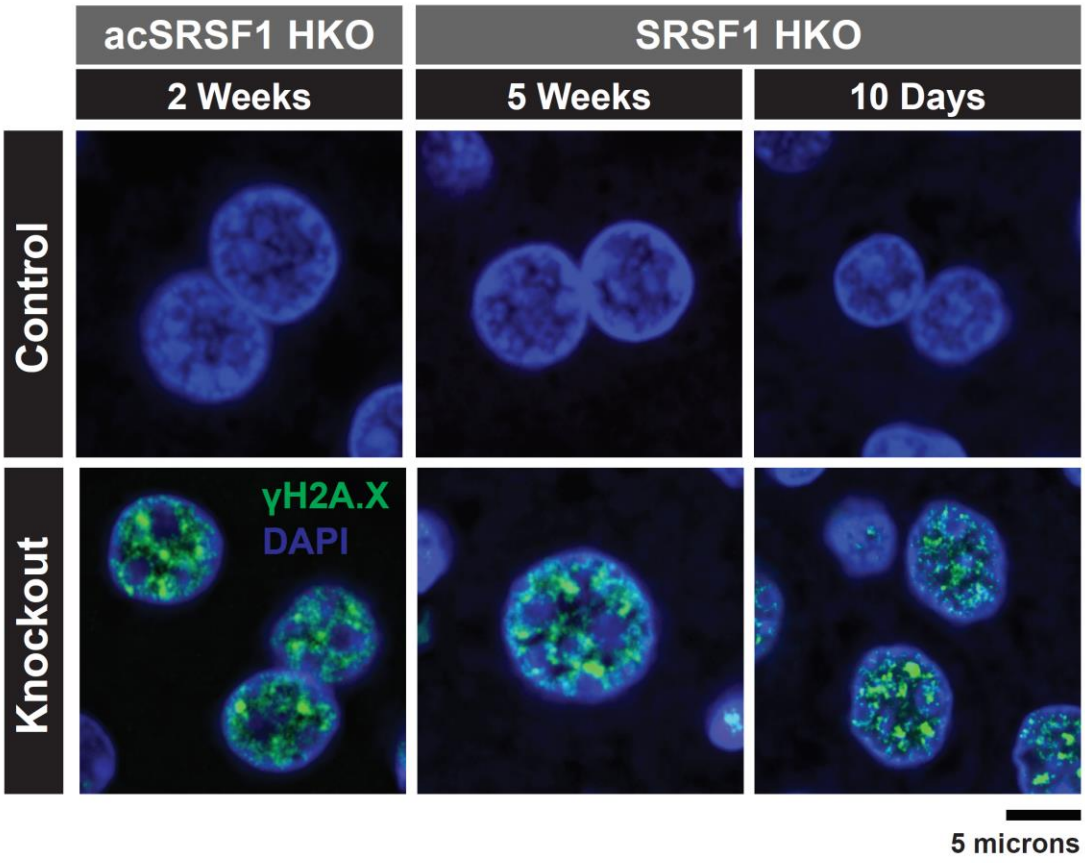


Figure 5.11: ***Loss of SRSF1 Activity Results in Immediate DNA Damage***

IF imaging of the DNA damage marker, γH2A.x, on acSRSF1 HKO and SRSF1 HKO liver tissue sections at the indicated timepoints. Images show representative hepatocyte nuclei among multiple samples per group (n = 4-5 per group). Loss of SRSF1 results in robust induction of widespread DNA damage.

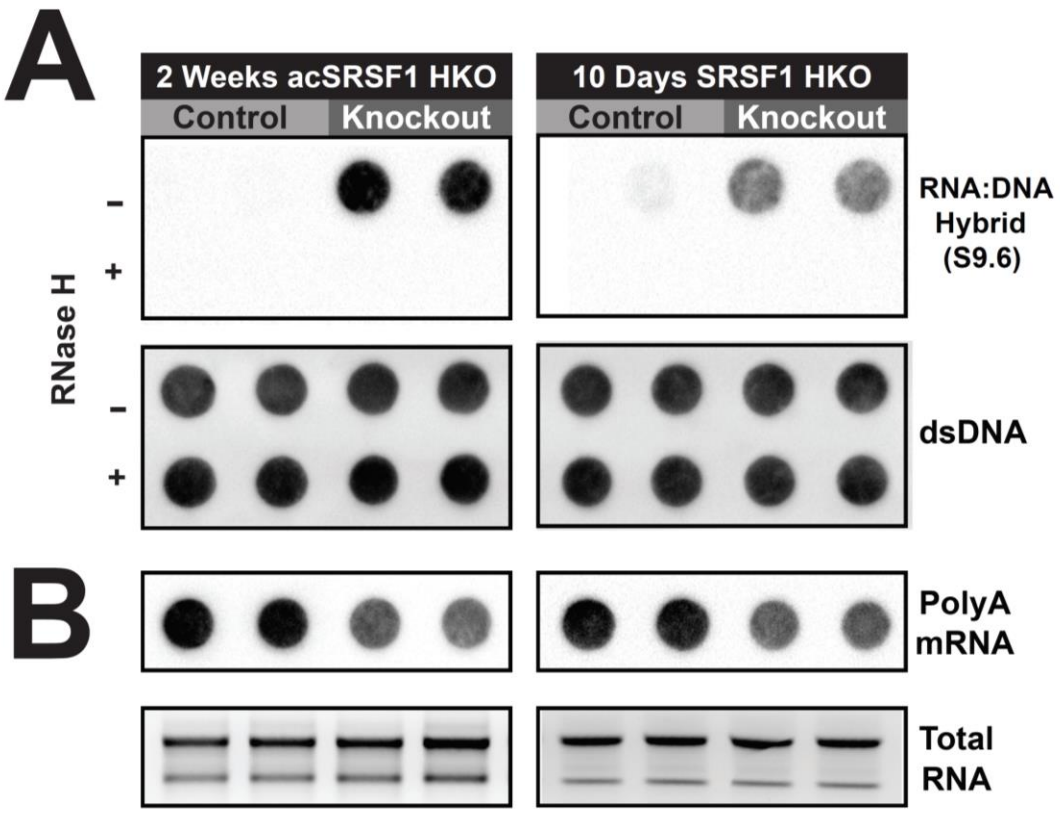


Figure 5.12: ***Loss of SRSF1 Leads to Accumulation of R-Loops***

(A) Representative blots of RNA-DNA hybrid dot blot assays using the S9,6 antibody on purified DNA from control and knockout samples from 2-week acSRSF1 HKO and 10-day SRSF1 HKO (n = 4 per group). Corresponding dot blot for dsDNA was performed as loading control. Presence of R-loop is confirmed with RNAse H treatment which results in the cleavage of R-loops and loss of signal. Striking signal is seen in knockout samples confirming accumulation of R-loops. (B) PolyA mRNA on 500 ng of isolated total RNA using a polydT probe. Total RNA gel electrophoresis image is shown for confirmation of equal loading. Knockout samples show decrease in PolyA mRNA signal suggesting decreased transcription.

Chapter 6: SRSF1 Knockout Leads to Global Translation Defect and Necroptosis

6.1 Introduction

It is now evident that loss of SRSF1 activity in hepatocytes results in widespread accumulation of R-loops and subsequent DNA damage. This activity of SRSF1 is a vital function in general physiology and is essential in protecting the genome from R-loop induced DNA damage. However, it is still not yet clear what are the downstream effects and responses to the widespread DNA damage. To gain further insight into the downstream effects, a global proteomics study was performed on hepatocytes isolated from acSRSF1 HKO. Understanding the changes occurring to the hepatocyte proteome will provide an overall picture of the mechanisms resulting in the eventual cell death.

With regards to understanding the eventual demise of the hepatocyte, previous investigations suggest the cells do not undergo apoptosis. This is apparent from the negative TUNEL staining. Histological appearance of the cell suggests the hepatocytes may likely be undergoing necrosis. To further investigate this possibility, western blot analysis for apoptosis and necrosis factors was performed.

6.2 Material and Methods

6.2.1 Global Proteomics Analysis of Hepatics Proteins by Mass Spectrometry

Samples were processed for proteomics according to the FASP protocol and desalted on an in-house prepared C18 tips (Wiśniewski et al. 2009). The peptides were separated on an EASY Spray C18 column (50 cm x 75 cm, 2 µm particle size) (ThermoFisher Scientific, Toronto) using an EASY nLC-1200. The mobile phase was composed of 0.1% formic acid in water (A) and 90% acetonitrile with 0.1% formic acid (B). The gradient was as follows: 5 - 40% B (0 – 120 minutes), 40% - 100% B (120 – 125 minutes), 100% B (125 – 135 minutes). The peptides were analyzed on a Thermo Q Exactive HF mass spectrometer in a Top 20 data dependent acquisition mode. Proteins were identified by searching the mouse (UP000000589) proteomes from UniProtKB (February 2019). Peptide spectral matches, and protein False Discovery Rates were set to 1%, and requiring a minimum of 1 unique peptide for identification. To increase the number of identified matches, match between runs" was enabled with a match time window of 0.7 minutes. Protein abundances were calculated using the iBAQ algorithm in MaxQuant.

6.2.2 Polysome Profiling of Isolated Hepatocytes from Acute SRSF1 HKO Mice

Polysome profiling was performed on acSRSF1 HKO mice at 2 weeks post injections of viral vectors. Isolated hepatocyte samples were prepared for both control (AAV8-TBG-GFP) and acSRSF1 HKO (AAV8-TBG-Cre) for polysome profiling following a previously described protocol (Seimetz et al. 2019). Hepatocytes were isolated in the same manner as described in *Section 9.1.2* with every buffer supplemented with 150 µg/mL cycloheximide. The pellets were flash frozen in liquid nitrogen and stored at -80

°C until ready for polysome fractionation. Frozen hepatocytes were thawed on ice for 15 minutes with 1 mL of polysome lysis buffer containing 10 mM Tris-HCl (pH 8.0), 150 mM NaCl, 5 mM $MgCl_2$, 1% Nonidet-P40, 40 mM dithiothreitol, 1 U/mL SUPERaseIn RNase inhibitor (Thermo Fisher) and 150 µg/mL cycloheximide. Thawed cells were pipetted gently 10 times to ensure lysis of cytoplasm. The cell nuclei and debris were removed by centrifugation at 12,000 x g for 1 minute at 4 °C. The supernatant was transferred to a fresh tube and then centrifuged again at 16,000 x g for 7.5 minutes at 4 °C to remove remaining cell debris and organelles. The resulting supernatant was transferred to a fresh tube and about 400 µL supernatant was layered onto a 12 mL linear sucrose gradient (10 – 50% sucrose (w/v) made using a Biocomp Gradient Master) and centrifuged in an SW-41Ti rotor (Beckman) for 125 minutes at 38,000 r.p.m. at 4 °C. The fractionated sample was gently moved through the detector using a peristaltic pump set at 3.5 mL/min with 60% sucrose as the chase solution. Polysome profiles were measured with a UA-6 absorbance (ISCO) detector at 254 nm and recorded using the associated Peak Chart software.

6.2.3 Global Translation Quantification using the SUnSET assay

Translating proteins were labeled using a protocol adapted from the Surface Sensing of Sensing of Translation (SUnSET) method (Goodman and Hornberger 2013; Goodman et al. 2010). The assay was performed on Control and acSRSF1 HKO mice at 2 weeks post injections of viral vector. Mice were injected with puromycin prepared in sterile PBS (0.04 µmol/gram body weight). After 45 minutes, livers were harvested, and protein lysates were prepared as described in *Section 9.1.3*. Proteins were separated by 10% SDS-PAGE. Puromycin-labeled peptides were identified using the mouse

monoclonal antibody 12D10 (EMD Millipore Catalog# MABE343). Protein synthesis levels were determined by densitometry analysis of whole lanes.

6.2.4 Serum Fractionation for Lipoprotein Particle Analysis

Plasma was collected from both control and acSRSF1 HKO mice at 2 and 4 weeks post viral injection as described in *Section 2.2.2* after a 6 hour fast starting from noon to 6 p.m. Plasma (~ 200 µL) was injected onto a Superose HR6 10/300 GL FPLC column (GE Healthcare). Lipoproteins were eluted with 24 mL of elution buffer containing 0.15 M NaCl, 1 mM EDTA, 0.2% w/v of sodium azide in PBS and 0.5 mL fractions were collected at a flow rate of 0.5 mL/min. Triglyceride and cholesterol concentration was measured using Infinity kits (Thermo Scientific). In a microtiter plate, 100 µL of plasma as well as standards were incubated with 100 µL of Infinity reagent and then incubated at 37 °C for 30 minutes. The plates were measured for absorbance using a BioTek machine at 500 nm. Using absorbance measurements, the quantity of triglyceride and cholesterol were calculated.

6.3 Results

6.3.1 Global Mass Spectrometry Analysis Reveals Depletion of Ribosomal Proteins

Global protein abundance estimates by mass spectrometry was able to detect a total of 3,603 proteins. Performing differential abundance analysis between control and acSRSF1 HKO hepatocytes reveals numerous proteins changing in abundances. Specifically, 613 different proteins were detected to be decreasing in abundance while 159 were found increasing **(Figure 6.1)**. It is not surprising that a greater proportion of proteins are decreasing in abundance since to cells is most likely shutting down translation in response to DNA damage.

Intersection of the proteomics data was performed with the available transcriptome data obtained for the acSRSF1 HKO model. Specifically, intersection was performed with corresponding genes with SRSF1 binding, gene expression changes, and splicing changes **(Figure 6.2 A-D)**. The intersection revealed expected findings such as genes found to be downregulated or upregulated in mRNA abundance also showed abundance changes in the same direction. No significant effects on abundance was found for genes exhibiting SRSF1 binding or differential splicing events.

One of the early concerns regarding the SRSF1 HKO model was the possibility of induction of other SR proteins to compensate the deficiency of SRSF1. This was not feasible as it would require obtaining the antibodies for the other 11 members of the SR family. However, this question can easily be answered by the estimations provided by the proteomics data. Estimations of the other SR protein factors show that except for SRSF4, all other SR protein abundances are stable and do not change between control and SRSF1 knockout **(Figure 6.3)**. SRSF1 was of course found to be significantly

downregulated as expected and SRSF4 was slightly increased in abundance. Interestingly, SRSF3, which has been shown to be involved in hepatocyte differentiation, has the highest abundance of all the SR proteins (Sen, Jumaa, and Webster 2013).

To gain an unbiased view of the various processes the differentially abundant proteins are involved with, a gene ontology analysis was performed. Results from the analysis show that the downregulated proteins are strongly enriched for ribosomal proteins and factors involved in translation **(Figure 6.4)**. Proteins that are increasing in abundance are enriched for factors involved in lipoprotein and cholesterol metabolism. To further understand the extent of downregulation of ribosomal proteins in acSRSF1 HKO, a heatmap was generated depicting the abundance changes of all detected proteins involved in translation **(Figure 6.5)**. Surprisingly, a striking and pervasive downregulation of ribosomal proteins, translation initiation factors, and tRNA synthetases is found at the level of protein abundance. However, mitochondrial ribosomal proteins showed an increase.

The remarkable downregulation of ribosomal proteins seen from the proteomics data suggests that acSRSF1 HKO mice will exhibit impaired translation. It is also known that Non-sense mediated decay (NMD) is dependent on translation. Therefore, impairment in translation activity will result in impairment of NMD. Impaired NMD would lead to increased abundance of intron containing NMD transcripts. To investigate this possibility, a global intron retention analysis was performed on the acSRSF1 HKO RNA-seq data **(Figure 6.6)**. Indeed, results from the intron retention analysis show that there is a robust increase in the levels of intron containing transcripts. This widespread

detection of intron containing transcripts further supports the notion that NMD is impaired rather than misregulation of SRSF1 mediated splicing of these transcripts.

6.3.2 Polysome Profiling of acSRSF1 HKO Hepatocytes Shows Absence of Polysomes

Proteomics and transcriptomics data from acSRSF1 HKO mice suggest there may be a defect in translation upon loss of SRSF1. To test this hypothesis, polysome profiling was performed since it can allow for the determination of ribosomal occupancy changes between two conditions. One possibility is that certain transcripts may exhibit lower translation upon knock-out of SRSF1. It is known that SRSF1 can regulate the translation of target mRNAs. Since polysome profiling results in the separation of mRNAs based on its ribosomal occupancy, differences in occupancy can be measured between Control and acSRSF1 HKO. Separation is performed on a sucrose gradient and mRNAs with increasing ribosomal occupancy appear as peaks on the chromatogram.

Surprisingly, the polysome profiles of acSRSF1 HKO show a near complete loss of polysomes **(Figure 6.7)**. While a monosome peak is present in the acSRSF1 HKO, subsequent peaks corresponding to polysomes are not which can be seen in Controls. This result agrees with the previous proteomics and transcriptomics data. Loss of higher polysomes suggests loss of SRSF1 results in a severe defect in translation. Interestingly, the profile for acSRSF1 HKO show the presence of a monosome peak. It is not clear if these are functional ribosomes translating the corresponding mRNA or if these are stalled ribosomes on the mRNA. To answer this question a different approach is required. While there is a correlation between ribosomal occupancy and translation rates, this technique does not allow for direct measurement of protein synthesis.

6.3.3 Global Translation is Diminished in acSRSF1 HKO Hepatocytes

Polysome profiling revealed absence of polysomes in acSRSF1 HKO. While this would suggest that global translation is impaired in acSRSF1 HKO, it does not provide definitive proof. To investigate the rate of global protein synthesis in the acSRSF1 HKO model, a SUnSET assay was performed. This assay is a pulse-chase experiment based on incorporation of puromycin to nascently synthesized proteins. Proteins are then isolated from the tissue and then western blot analysis is performed using puromycin antibody. The resulting signal is normalized to total protein signals to determine the rate of global protein synthesis. The assay was performed on acSRSF1 HKO mice at the 2-week timepoint **(Figure 6.8 A)**. In addition to the liver, spleen was collected as a normal tissue control from the acSRSF1 HKO model. Quantification of the puromycin incorporation shows that acSRSF1 HKO have severely diminished protein synthesis activity **(Figure 6.8 B)**. This provides strong proof that global translation in acSRSF1 HKO mice is diminished.

6.3.4 Lipoprotein Particle Formation is Depleted in SRSF1 HKO Mice

Due to the dramatic decrease in global protein synthesis in acSRSF1 HKO hepatocytes, these hepatocytes are expected to display deficiency in their overall functions. One of the major functions of a hepatocyte is the processing of lipids into lipoprotein particles and release into circulation for systemic delivery. Hepatocytes generate most high-density lipoprotein particles that are present in circulation. To determine if acSRSF1 HKO have impairment in lipoprotein particle formation, fractionation was performed on serum collected from fasted mice at the 2- and 4-week

timepoint **(Figure 6.9)**. As expected, serum at the 4-week timepoint are depleted of HDL particles suggesting that liver is failing to synthesize the particles.

6.3.5 Western blot Analysis of p53, eIF2α, and Cell Death Factors

It is evident that SRSF1 knockout results in striking impairment of global translation. However, the mechanism resulting in the global shutdown of translation is not yet clear. One possible explanation is that the widespread DNA damage is activating p53 which is known to subsequently result in the global shutdown of translation. This response is thought to protect the cell from generating aberrant protein and to focus efforts into repairing the damaged DNA. To this investigate this hypothesis, western blot analysis was performed for total p53 and phospho-p53 (S15) which is the active form **(Figure 6.10)**. Surprisingly, p53 activation is not seen by western blot analysis in acSRSF1 HKO timepoints. It is also known that phosphorylation of eIF2α signals shutdown of global translation. It is possible that DNA damage could lead to global translation inhibition via phosphorylation of eIF2α. However, like p53, eIF2α was not found to be deactivated in acSRSF1 HKO **(Figure 6.10)**.

Western blot analysis was also performed to measure BAX and RIPK1 proteins, which are classical apoptosis and necrosis factors, respectively. Interestingly, both BAX and RIPK1 showed robust induction at the 4-week timepoint in acSRSF1 HKO model **(Figure 6.10)**. This suggests that the cell death occurring in this model is via necroptosis, explaining why the TUNEL staining are negative. Due to the impairment of global translation it is likely that the SRSF1-deficient cell is not able to trigger the apoptosis pathway resulting in eventual necrosis.

6.4 Conclusion

Global proteomics data showed a striking downregulation of multiple factors involved in the initiation of translation and ribosomal proteins **(Figure 6.5)**. It was hypothesized that diminished levels of these factors would lead to impairment of global translation. To test this hypothesis, both polysome profiling and SUnSET assays were performed on acSRSF1 HKO hepatocytes at the two-week timepoint. Polysome profiling allows for the isolation and separation of mRNAs based on the number of bound ribosomes. Transcripts which are translated at a higher efficiency tend to have a greater number of bound ribosomes. Indeed, results from this assay showed that acSRSF1 mice had an absence of polysomes **(Figure 6.7)**. While this assay provides insight of the polysome distribution in the cell, it does not give direct estimate of translation rates or protein synthesis.

To address this question, the SUnSET assay, a nonradioactive method to monitor protein synthesis, was utilized. This technique employs the use of the antibiotic puromycin, a structural homolog of tyrosyl-tRNA, which can be readily incorporated into nascent peptide chains. Using an anti-puromycin antibody, the amount of puromycin incorporation can be quantified with western blotting, hence, providing an indirect measure of protein synthesis. In agreement with the proteomics and polysome profiling data, acSRSF1 HKO exhibited severely diminished rates of global protein synthesis **(Figure 6.8)**. Taken together, loss of SRSF1 activity results in eventual cessation of global translation. However, it is unclear what triggers this shutdown of translation.

To further investigate possible mechanism of translation shutdown, western blot analysis was utilized to probe levels of known factors involved in translation inhibition. A

possible mechanism resulting in translation inhibition is activation of p53 since DNA damage is present in acSRSF1 HKO. It is known that activated p53 triggers shutdown of global translation. Surprisingly, p53 was not activated in the by western blot analysis. There is now upcoming evidence showing that there are conditions in which DNA damage results in a p53-indpendnet response (Fagan-Solis et al. 2020; Riepe et al. 2018). Another possible explanation of the absence of p53 activation is that the activation occurs within a short window which was not captured at the timepoints which were collected. A second possible mechanism of global translation inhibition is phosphorylation of eIF2α. This event is a well-established mechanism for cause global translation inhibition. Like p53, phosphorylation of eIF2α was also absent in the acSRSF1 HKO. Currently, the mechanism resulting in the inhibition of translation is not clear and further investigation is required.

Western blot analysis was also performed for BAX and RIPK1 which are known apoptosis and necrosis factors, respectively. Interestingly, both factors were in the acSRFS1 HKO, indicating that the SRSF1-deficient hepatocytes undergo necroptosis **(Figure 6.10)**. This explains the negative TUNEL staining in the SRSF1 HKO mice. Due to the global impairment of translation, it is possible that factors necessary for the complete execution of apoptosis may be deficient. Therefore, the failed apoptosis might result in eventual necrosis as seen by the elevated levels of RIPK1.

6.5 Chapter Figures

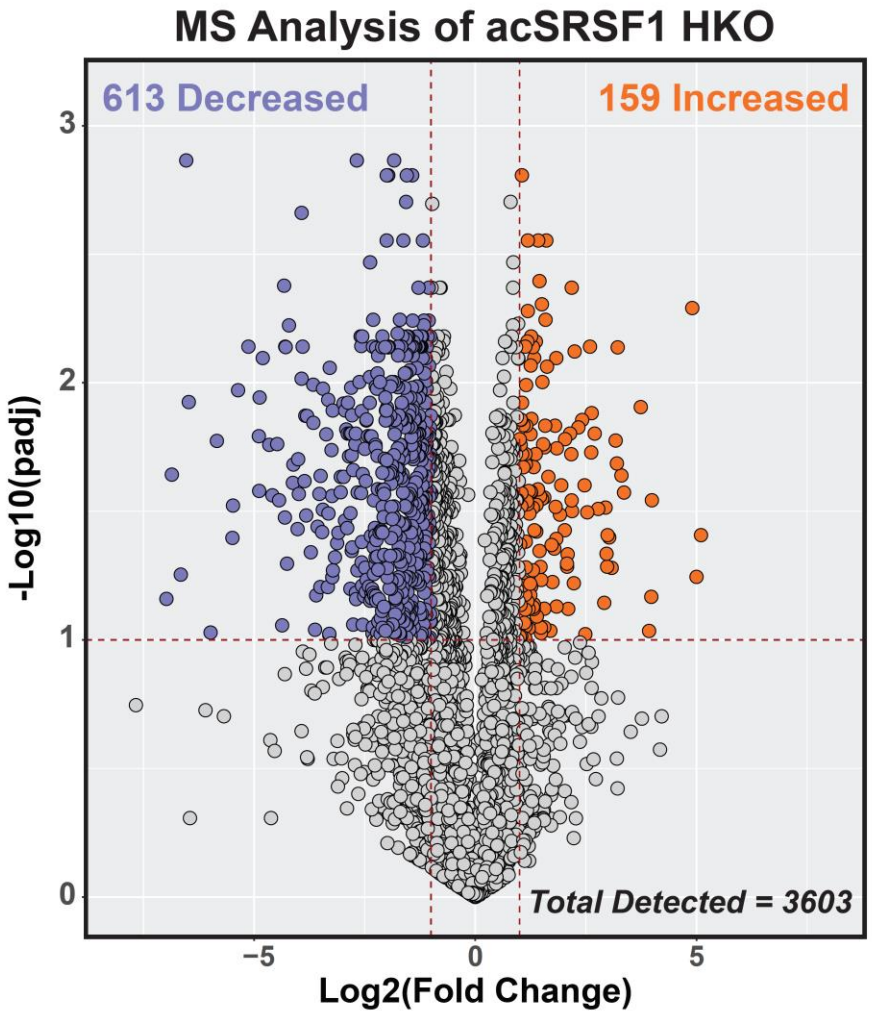


Figure 6.1: ***AcSRSF1 HKO Hepatocytes Exhibit Extensive Remodeling of the Proteome***

Volcano plot depicting the changes in protein abundances estimated from global proteome analysis on hepatocytes isolated from control and acSRSF1 HKO mice using mass spectrometry (n = 3 per group). A total of 3,603 proteins were detected from mass spectrometry. Of the 772 proteins identified to be significantly changing in abundance (Log2 Fold Change > 1 and FDR < 0.1), 613 (~79%) were found to be downregulated. Differential abundance was determined by calculating fold change ratios using IBAQ quantification values.

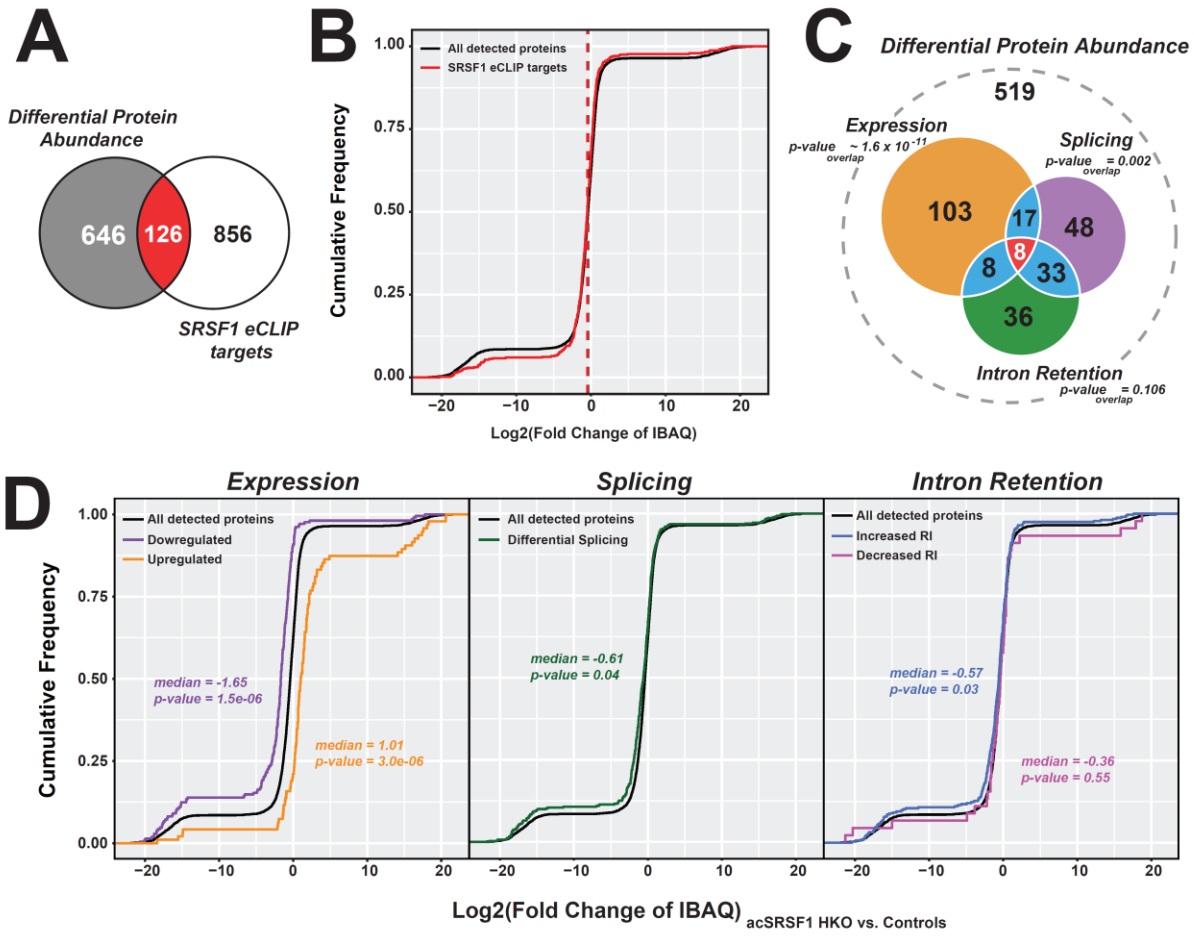


Figure 6.2: ***Majority of Proteome Remodeling in acSRSF1 HKO are Independent of the Transcriptome***

(A) Venn diagram showing overlap of proteins with abundance changes in acSRSF1 HKO model and associated genes with SRSF1 binding. A total of 126 genes with SRSF1 binding showed changes in abundance of the associated protein. (B) Cumulative plot of Log2 Fold Change of protein abundances with SRSF1 binding in associated genes. It appears protein abundances are independent of SRSF1 binding. (C) Venn diagram of protein abundance overlap with genes changing in expression, splicing, and intron retention. Significant overlap is seen with genes changing in expression and splicing. (D) Cumulative frequency plot of Log2 Fold Change of protein abundance for indicated subset.

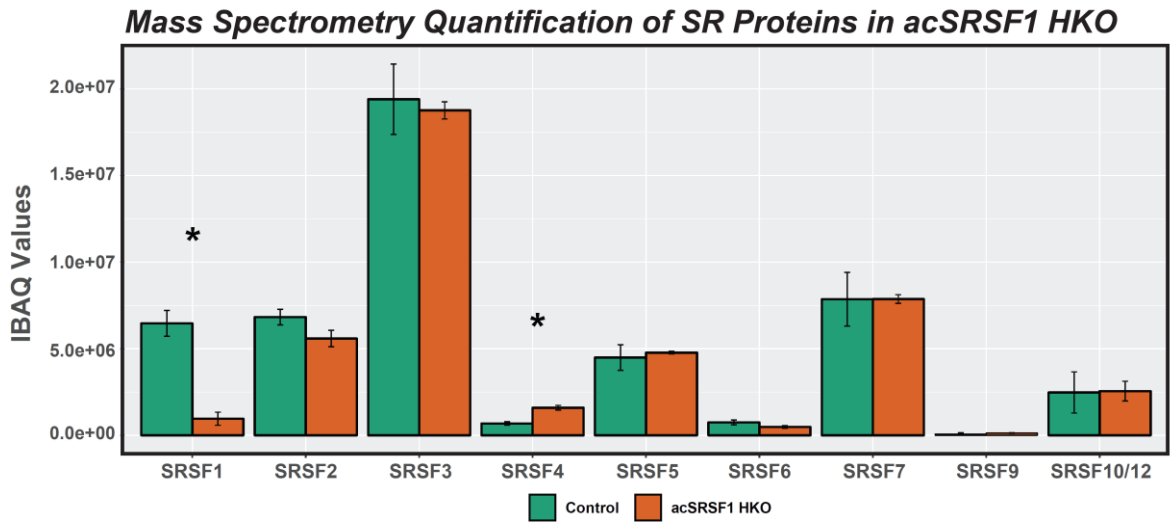


Figure 6.3: ***Abundances of Other SR Proteins are not Changing upon SRSF1 Knockout***

Plot of IBAQ quantification values for the detected SR proteins in acSRSF1 HKO mass spectrometry analysis. In agreement with the model, SRSF1 is detected to be significantly diminished in abundance in the knockout model. Of all the SR proteins detected, only SRSF4 shows slight elevation in abundance upon knockout of SRSF1.

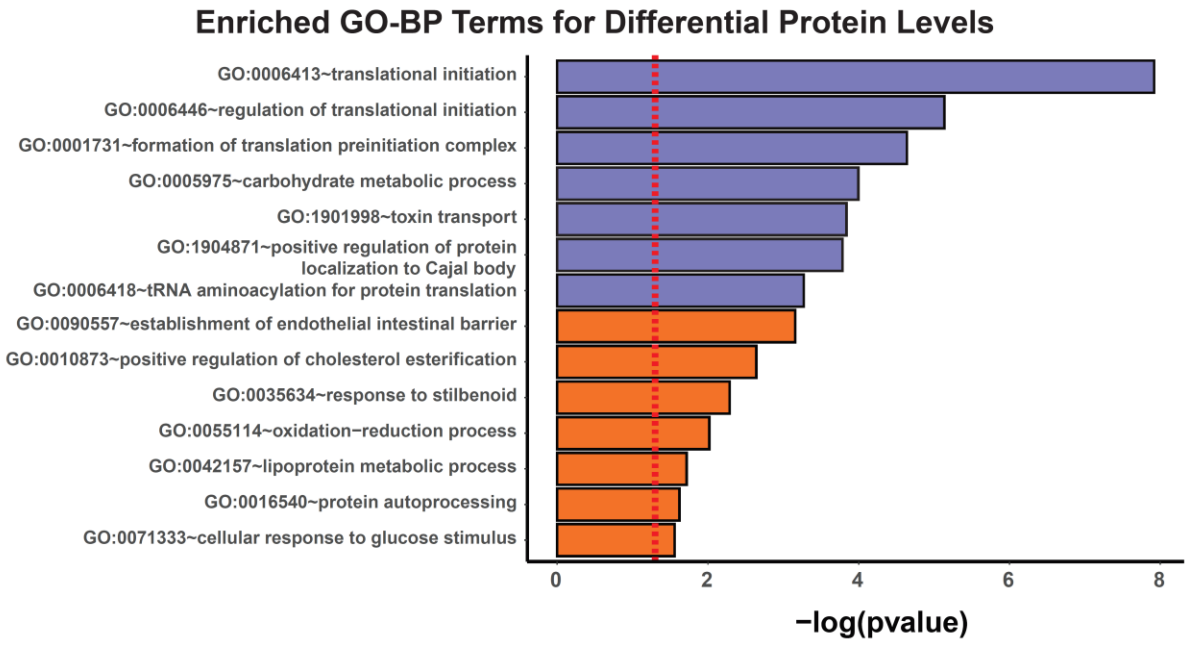


Figure 6.4: ***Gene Ontology Analysis of Proteins Exhibiting Differential Abundances***

Gene ontology analysis of proteins with differential abundances in acSRSF1 HKO. GO term enrichment analysis was performed on downregulated and upregulated proteins separately. Proteins that were downregulated in acSRSF1 HKO have a strong enrichment for proteins involved in translation.

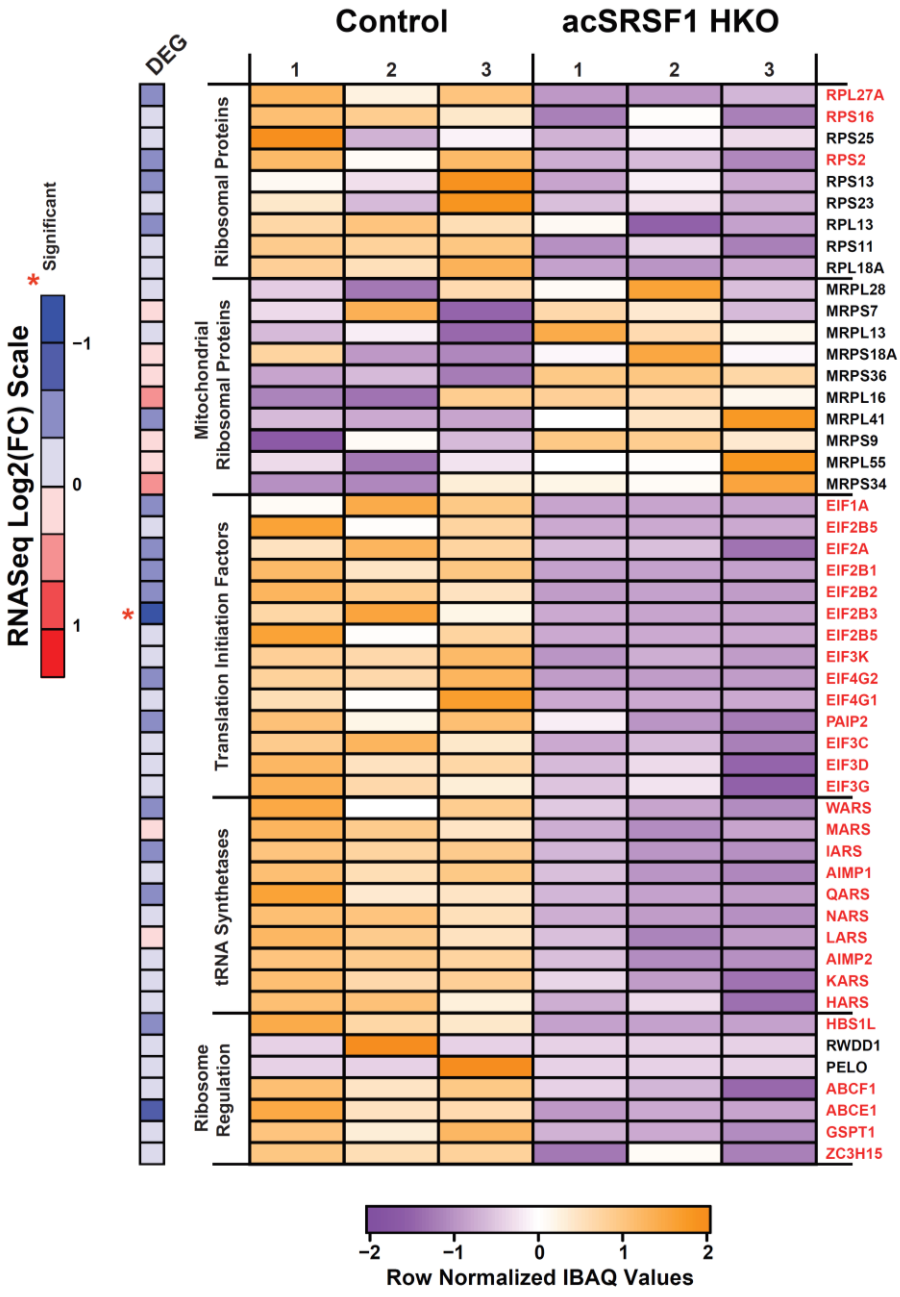


Figure 6.5: ***Hepatic Knockout of SRSF1 Results in a Striking Downregulation of Factors Involved in Translation***

Heatmap showing row normalized IBAQ quantification values for detected proteins involved in translation regulation. The heatmap is divided into subsections involved in different aspects of translation. Red colored labels signify proteins which are significantly changing in protein abundance. The heatmap bar on the left shows the RNA-Seq differential log2 fold change of the associated gene.

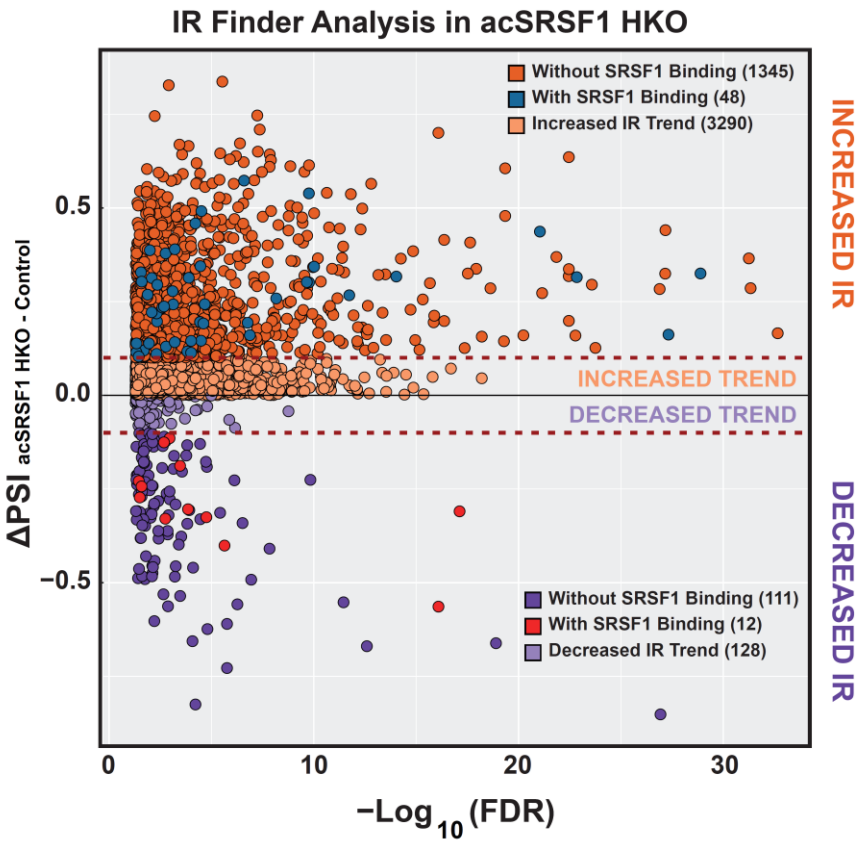


Figure 6.6: ***AcSRSF1 HKO Exhibit Increased Intron Retention Events***

Volcano plot of intron retention events identified from the acSRSF1 RNA-seq. A total of 1,516 introns were determined to be differentially retained in the acSRSF1 HKO. Of the differentially retained introns, over 90% exhibited increased retention. Furthermore, a small fraction with about 4% of retained intron events had SRSF1 binding.

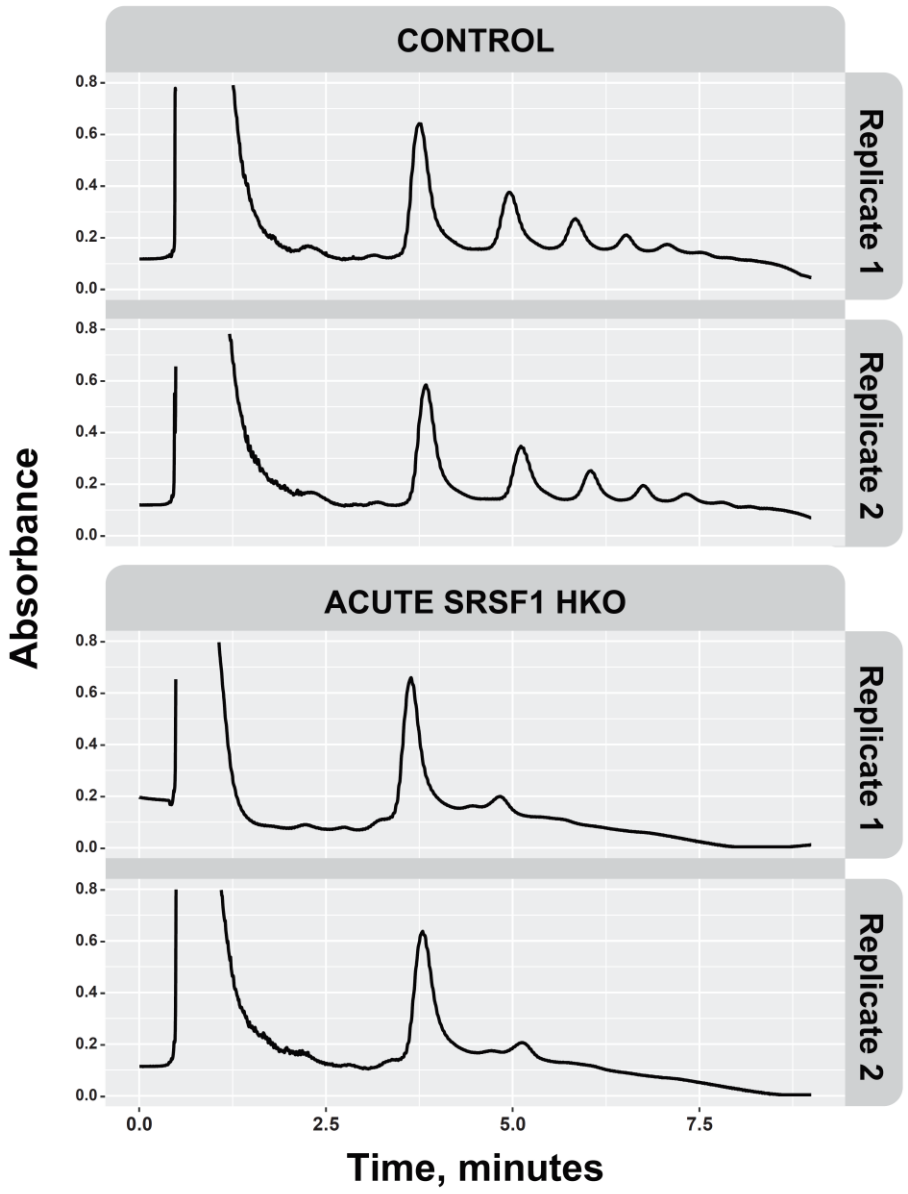


Figure 6.7: ***Polysome Profiling Reveals Absence of Polysomes in acSRSF1 HKO***

Polysome profiling performed on control and acSRSF1 HKO hepatocytes. In control samples, the monosome peak appears at ~4 minutes followed by multiple polysomal peaks. In acSRSF1 HKO, the initial peak is left shifted in comparison to controls which may signify incomplete ribosomal assembly on transcripts. Furthermore, polysome peaks are absent in acSRSF1 HKO samples.

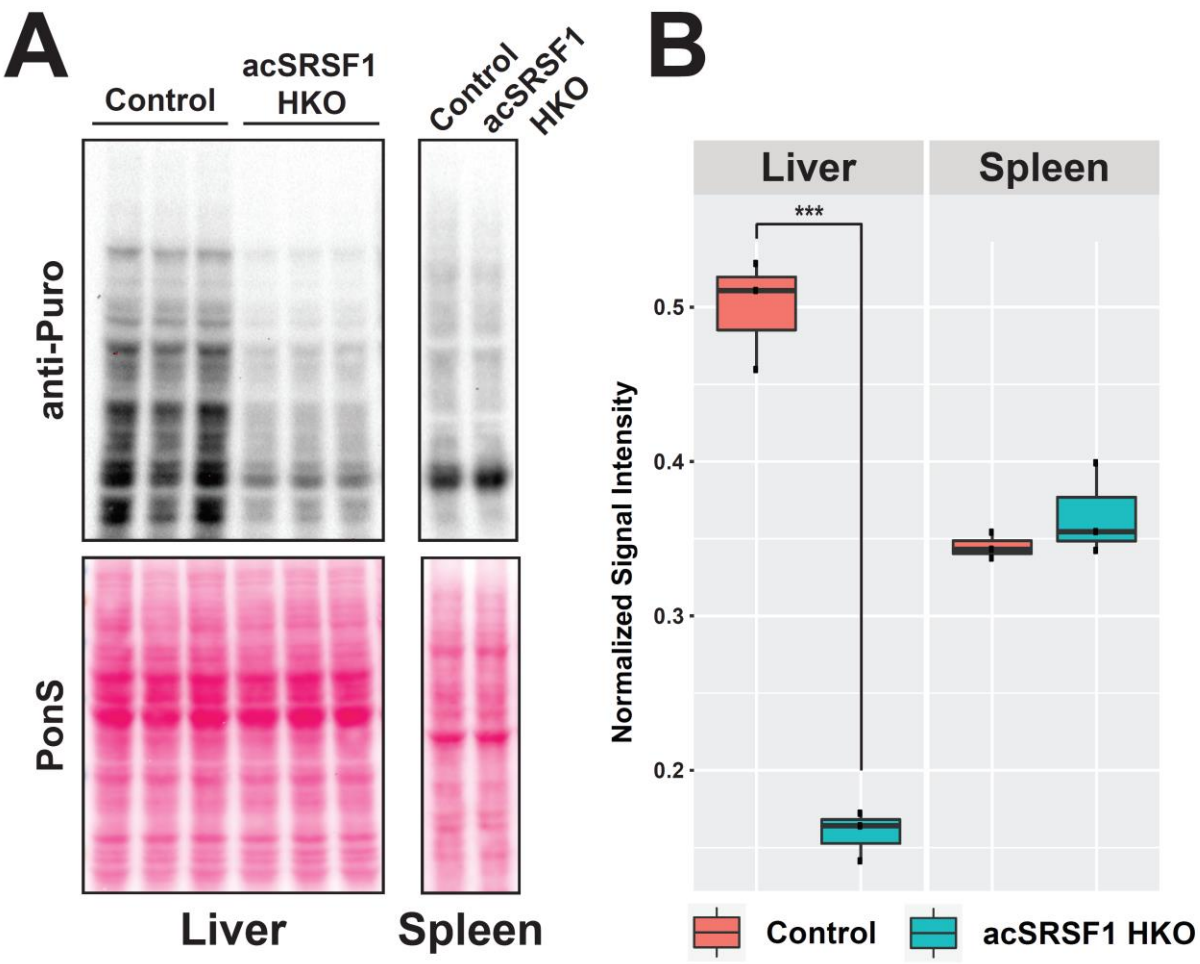


Figure 6.8: ***Global Translation is Strikingly Diminished in acSRSF1 HKO Hepatocytes***

(A) Puromycin incorporation assay was performed in hepatocytes isolated from acSRSF1 HKO mice 2-weeks post viral induction to determine global protein synthesis. Assay was performed on spleens as a control tissue. (B) Quantification of nascent protein synthesis by normalizing lane signal of anti-Puromycin blot by the Ponceau S signal. The acSRSF1 HKO mice exhibit drastic reduction in protein synthesis in comparison to control.

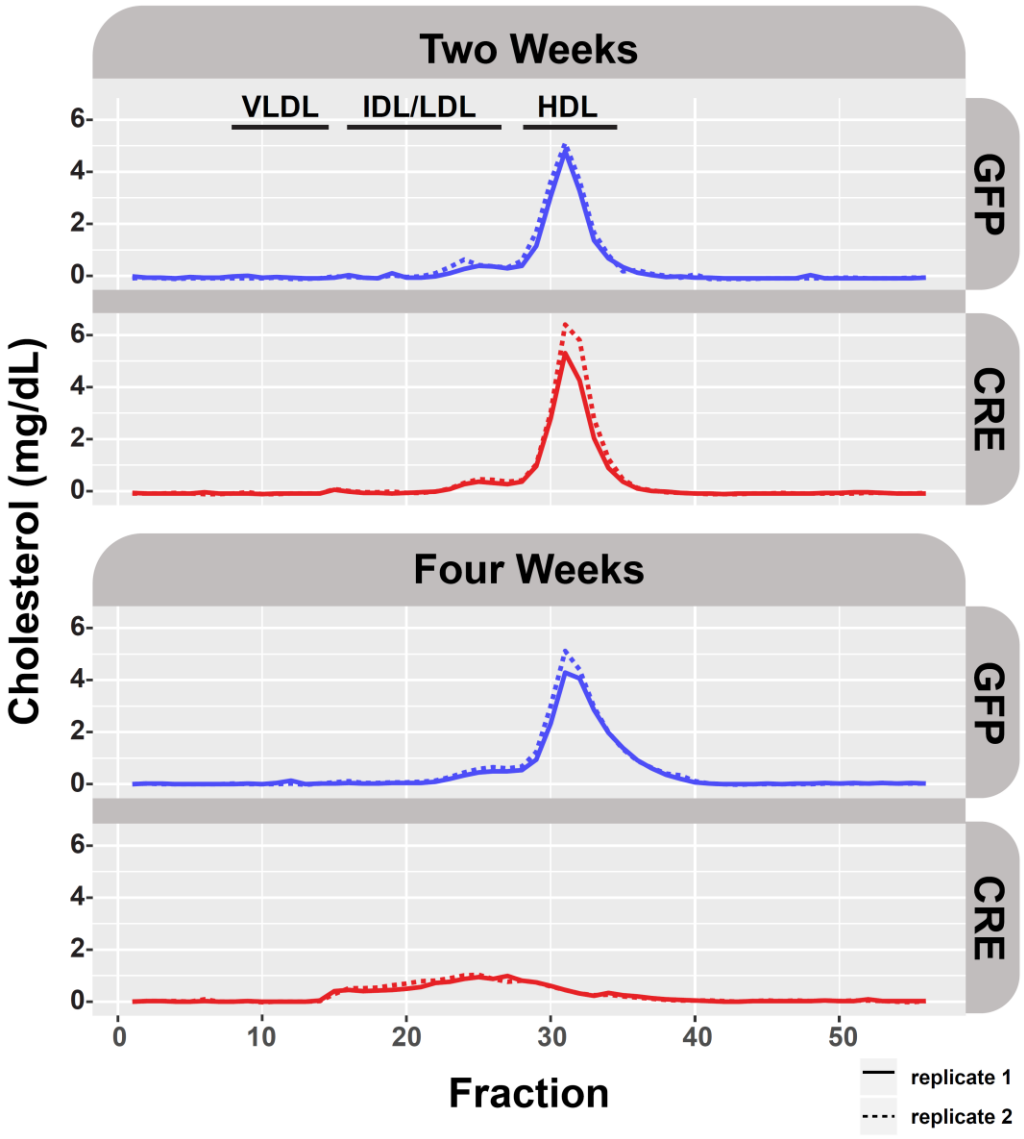


Figure 6.9: ***Serum Lipoprotein Particle Fractionation Reveals Diminished Particles in acSRSF1 HKO***

Lipoprotein particle fractionation was performed on serum collected from fasted control or acSRSF1 HKO mice at the indicated timepoints post viral induction. Each replicate consists of serum pooled from 2 biological samples. HDL particles are severely diminished in the 4-week acSRSF1 HKO mice serum. This suggests that SRSF1 knockout in hepatocytes results in loss of hepatic functions.

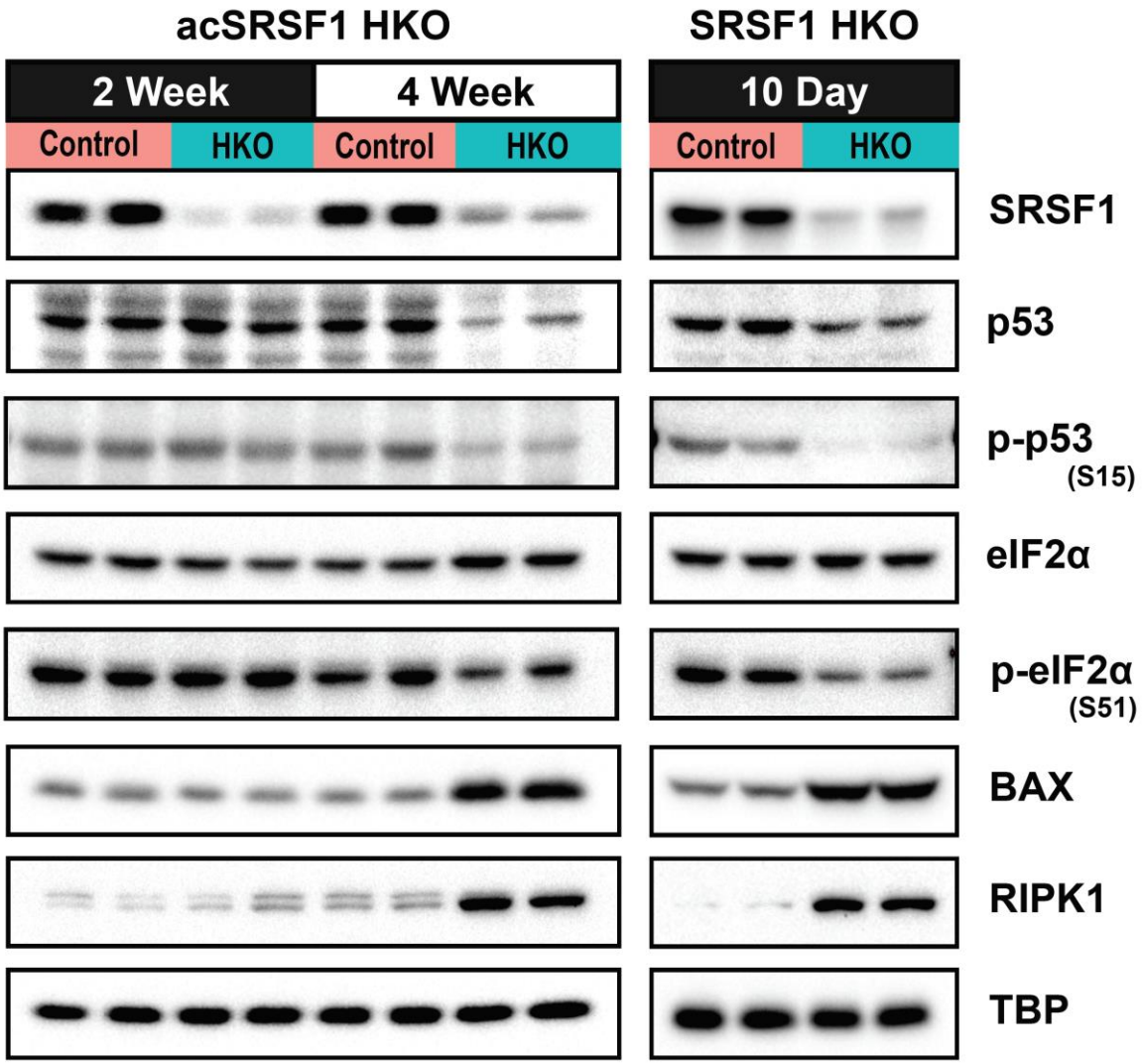


Figure 6.10: ***Hepatocyte deficient of SRSF1 do not trigger Phosphorylation of p53 or eIF2α and Eventually Undergo Necroptosis***

Western blot analysis of indicated factors at the 2- and 4-week acSRSF1 HKO and 10-day SRSF1 HKO models. Blots depict representative samples for each group (n = 4 – 5 per group). Levels of p53 are decreased at the 4-week timepoint in acSRSF1 HKO and 10-day SRSF1 HKO model. Both models exhibit evidence of necroptosis with increased levels of BAX and RIPK1 at 4-week in acSRSF1 HKO and 10-day in SRSF1 HKO.

Chapter 7: Loss of SRSF1 Activity in HepG2 Cells Leads to DNA Damage and Protein Synthesis Impairment

7.1 Introduction

Ablation of SRSF1 in hepatocytes of mouse liver results in immediate death of the cell and development of pathology. Molecular investigations of the SRSF1 HKO models has revealed widespread DNA damage which is most likely induced by pathological accumulation of R-loops. This is further supported by transcriptomics data which shows induction of DNA damage response and repair genes. Furthermore, proteomic studies of the SRSF1 HKO models show severe depletion of ribosomal proteins which result in global impairment of translation.

To determine if this misregulation occurs in the context of human cells, publicly available ENCODE data was analyzed. The ENCODE project performed a large study in which various RNA binding proteins were knocked down in HepG2 cells and subsequent RNA-Seq was performed. This dataset includes knockdown of various SR proteins. Transcriptome analysis of this dataset was compared to the results of the acute SRSF1 HKO to determine which aspects are conserved. In addition to SRSF1, other SR proteins such as SRSF3, SRSF5, SRSF7 and SRSF9 knockdowns have also been performed by the ENCODE consortium. These datasets have also been analyzed to find similarities and differences between knockdown of different SR proteins.

Furthermore, knockdown experiments of SRSF1 in HepG2 cells was performed to first confirm if DNA damage and translation defects can be replicated in a cell culture model and secondly to further explore the mechanism of the translation defect.

7.2 Material and Methods

7.2.1 Analysis of SR Protein Knockdown ENCODE Dataset

Analysis of publicly available dataset were performed using the same pipeline described in *Section 3.2.1*. Briefly, differential gene expression analysis was performed using kallisto, tximport and DESeq2 packages. Differential splicing analysis was performed using rMATS. Gene ontology analysis was performed using the gProfiler web-based tool. Details regarding the datasets used for this analysis can be found in *Section 9.2*. Briefly, ENCODE datasets of SRSF1, SRSF3, SRSF5, SRSF7 and SRSF9 protein knockdown by shRNA in HepG2 cells was used for the study (Sundararaman et al. 2016).

7.2.2 Culturing of HepG2 Cells and siRNA Mediated Knockdown

HepG2 cell line was obtained from ATCC (catalog HB-8065) and cultured according to ATCC specifications. Cells were cultured in DMEM supplemented with 10% FBS, 2 mM glutamine, and 10 U/ml penicillin and streptomycin. For knockdown experiments silencer select siRNAs against *SRSF1* and *TP53* (Thermo Fisher Scientific #4392420 and #4390824) along with a negative control (Thermo Fisher Scientific #4390843) were purchased. Approximately 500,00 cells were seeded into a 6-well format and were reverse transfected with 20 nM of gene-specific siRNA oligos using RNAiMax (Thermo Fisher Scientific #13778075) and then transfected again using forward transfection 36 hours with 20 nM of siRNA. Cells were harvested after 72 hours starting from the initial reverse transfection. For end-point assays requiring fluorescent imaging, cells were grown on coverslips coated with poly-L-lysine, 0.1% (w/v) (Sigma-Aldrich #P8920).

Briefly for immunofluorescent staining, cells were fixed for 10 minutes using 4% PFA solution followed by a PBS wash three times and then permeabilized with 0.2% Triton X-100 plus 1% normal goat serum (NGS) in PBS/pH 7.3 for 5 minutes on ice. Cells were then washed with PBS with 1% NGS and then incubated in primary antibody dilution for 1 hour at room temperature in a humidified chamber. This was followed by another wash and then incubation with secondary antibody dilution for 1 hour at room temperature. The cells were then washed with PBS and stained with DAPI before inverting the coverslip onto a glass slide with aqueous mounting media. For details on the dilutions of antibody used, refer to *Section 9.1.9*.

Additional fluorescent assays performed on the knockdown HepG2 cells included a fluorescent protein synthesis assay and Annexin V staining. For the fluorescent protein synthesis assay, the Click-iT™ HPG Alexa Fluor™ 488 Protein Synthesis Assay Kit (Thermo Scientific #C10428) was used following manufacturer's instructions. Briefly, 2 hours prior to harvesting, culture media was replaced with methionine-free DMEM media. Cells were incubated in media for 1 hour before supplementing well media with 50 µM of ClickIT-HPG reagent. Cells were incubated for additional 1 hour before fixation. Fixed cells were prepared following standard kit protocol. For the fluorescent Annexin V cell death staining, the Dead Cell Apoptosis Kit with Annexin V FITC and PI (Thermo Scientific #V13242) was used following manufactures instructions. For western blot analysis, cells were directly lysed in 2X Laemmli buffer and then prepared for PAGE gel electrophoresis as described in *Section 9.1.3*.

7.3 Results

7.3.1 Expression of SR Proteins in HepG2 Cell Line

Before starting analysis of the shRNA mediated knockdown of SR proteins in HepG2 cells, the expression levels of the SR genes in HepG2 cells were determined. This was achieved by plotting TPM data from all control shRNA knockdown HepG2 sample data available on ENCODE **(Figure 7.1)**. Expression of SR genes were at reasonable expression levels with SRSF5 exhibiting the highest expression levels in HepG2. Interestingly, SRSF3 was found to be the most abundant protein in mice hepatocytes.

7.3.2 Knockdown of SR Proteins Results in Significant Impact to the Transcriptome

Next, to understand the transcriptomics changings occurring in the SRSF1, 3, 5, 7, and 9 protein knockdowns in HepG2 datasets, differential expression and splicing analysis was performed. Analysis revealed a substantial number of changes in gene expression in splicing among the knockdowns **(Figure 7.2 and 7.3)**. Interestingly, of all the SR protein knockdowns, SRSF1 showed the largest number of changes in gene expression and splicing **(Figure 7.8)**. Spearman correlations were performed on the data to gain insight into which factors share most similarity with regards to the changes seen in the transcriptome **(Figure 7.4)**. Correlation analysis shows that SRSF1 has strongest correlation with SRSF7. Interestingly, of the SR proteins, SRSF7 shares the most similarity to SRSF1 and is also has shuttling activity like SRSF1.

Gene ontology analysis of the differentially expressed genes and spliced events was performed to further understand what processes are enriched in the various gene sets. In the upregulated gene sets for the various SR protein knockdowns, SRSF1 shows a strong enrichment for genes involved in cell cycle and mitosis **(Figure 7.5)**. Interestingly

SRSF7 enriched for processes involved in cell adhesion and cell motility. However, in the downregulated gene sets, SRSF1 and SRSF7 display a strong enrichment for genes involved in translation and ribosomal proteins **(Figure 7.6)**. Similar enrichment for these processes were seen in the acSRSF1 HKO downregulated proteins. This finding suggests that decrease levels of translation factors might be driven at the level of transcript abundance. Like the SRSF1 HKO, the differential splicing gene set does not exhibit strong enrichment for any biological process **(Figure 7.7)**.

Intron retention analysis was also performed on the shRNA mediated SR protein knockdown samples in HepG2. Interestingly, the SRSF1 knockdown dataset also displays a larger number of intron retention events with a greater portion of events resulting in increased retention **(Figure 7.8)**. This is comparable to what was seen in the acSRSF1 HKO model. It is possible this increased retention is a result of decreased translation activity as the downregulated genes strongly enrich for genes involved in translation.

7.3.3 Knockdown of SRSF1 Shows an Induction of a p53 Signature

Despite the widespread DNA damage observed in SRSF1 HKO mice, activation of p53 was not observed. This is unexpected since p53 is known to increase in abundance and become activated in response to DNA damage. Two possible explanations for this finding are that 1) the duration of p53 activation is short lived and was therefore missed within the timeframes the samples were collected or 2) DNA damage induced by R-loops is not detected by the complex that signals to p53. To investigate p53 status in the HepG2 SRSF1 knockdown dataset, expression of various p53 targets were examined (**Figure**

7.9). Surprisingly, multiple p53 targets are upregulated in the SRSF1 knockdown dataset suggesting that p53 is most likely activated.

7.3.4 Knockdown of SRSF1 in HepG2 Recapitulates SRSF1 HKO Phenotype

To further explore whether loss of SRSF1 in HepG2 does indeed result in DNA damage and inhibition of global protein synthesis, knockdown experiments using siRNA were performed in HepG2 cells. Efficient knockdown of SRSF1 was achieved using a combination of reverse transfection followed by a forward transfection after 24 hours **(Figure 7.10)**. Interestingly, knockdown of SRSF1 in HepG2 shows robust upregulation of p53 abundance. To examine DNA damage, immunofluorescent staining for γH2A.X was performed on the SRSF1 knockdown HepG2 cells. In agreement with the SRSF1 HKO findings, robust staining of γH2A.X is observed signifying DNA damage. Furthermore, a fluorescent based protein synthesis assay was also performed on these cells simultaneously. As expected, loss of SRSF1 results in decreased protein synthesis.

7.3.5 Protein Synthesis Inhibition in SRSF1-deficient Cells is Independent of p53

Next, to determine if the decrease in protein synthesis is caused by the activation of p53, a double knockdown experiment was performed where both SRSF1 and p53 were knocked down. Interestingly, co-knockdown of SRSF1 and p53 did not rescue the decreased protein synthesis phenotype **(Figure 7.11)**. This result signifies that the decrease in protein synthesis is independent of p53 activation. Finally, cells were cultured for 96 hours and then stained with Annexin V to probe for cells undergoing cell death. As expected, SRSF1 knockdown cells display robust staining with Annexin V **(Figure 7.11)**. Therefore, loss of SRSF1 activity in HepG2 also shows eventual death of the cell.

7.4 Conclusions

Knockdown of SRSF1 in HepG2 cells corroborates the findings observed in the SRSF1 HKO models. Interestingly, transcriptome analysis of the SRSF1 knockdown in HepG2 dataset showed a strong downregulation of ribosomal protein genes and genes involved in translation **(Figure 7.6)**. While SRSF1 HKO model did not show downregulation of genes involved in translation at the transcript level, mass spectrometry analysis displayed a striking downregulation in translation related factors. Furthermore, examining expression changes of p53 target genes, SRSF1 knockdown displayed a strong 53 activation signature **(Figure 7.9)**. Indeed, western blot analysis on SRSF1 knockdown HepG2 cells after 72 hours exhibited robust induction of p53 levels **(Figure 7.10)**. This was not observed in the SRSF1 HKO model. A possible explanation of this finding is that p53 activation is short lived and is therefore missed within the timeframes the hepatocytes were collected in the SRSF1 HKO models. Furthermore, levels of BAX and RIPK1 were not found to be induced at this timepoint. This suggests that the SRSF1 knockdown cells at this timepoint are likely accumulating DNA damage, however, cell death signals have not yet been triggered.

To confirm if DNA damage is indeed occurring after SRSF1 knockdown, immunofluorescent staining for γH2A.X was performed. Like the SRSF HKO, loss of SRSF1 in HepG2 shows presence of numerous γH2A.X foci in the nuclei confirming that DNA damage is present **(Figure 7.11)**. Next, to check if protein synthesis is inhibited a fluorescent based assay using HPG incorporation was performed which showed striking reduction of protein synthesis in SRSF1 knockdown **(Figure 7.11)**. While western blot analysis shows that cell death has not been triggered at the 72-hour timepoint, it is

expected that SRSF1 cells eventually undergo cell death. To investigate this possibility, SRSF1 knockdown was performed for 96 hours and then stained with Annexin V, a cell death marker. In agreement with expectation, longer knockdown of SRSF in HepG2 show prominent Annexin V signal. Therefore, SRSF1 knockdown in cells leads to eventual death of the cell **(Figure 7.11)**.

To check if the reduced protein synthesis is dependent on p53 activation, a co-knockdown of SRSF1 and p53 experiment was performed. Results from this experiment show that the rescue failed with persistence of the diminished protein synthesis **(Figure 7.11)**. Thus, despite the activation of p53, the reduced protein synthesis response is independent of p53. Currently, it is unclear the exact mechanism resulting in this phenotype. With the widespread R-loop accumulation and DNA damage, several probable mechanisms can be formulated to explain the reduction in global protein synthesis. However, testing the exact mechanism of his irregularity is outside the scope of the current study.

7.5 Chapter Figures

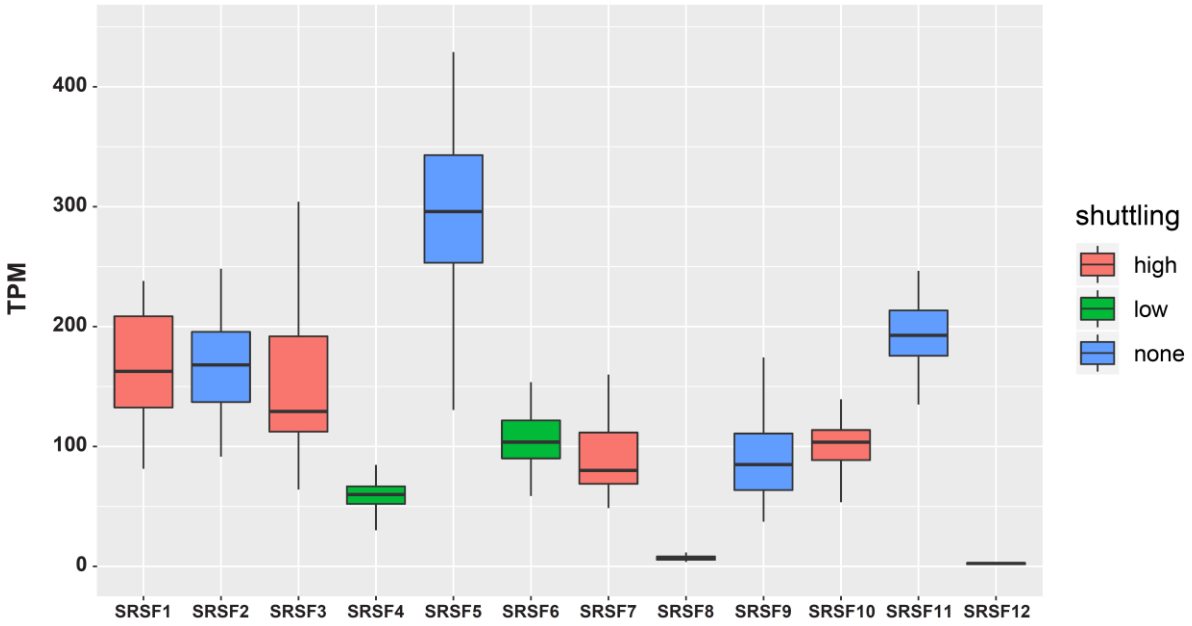


Figure 7.1: ***Expression of SR Proteins in HepG2 Cell Line***

Boxplot of TPM values for the various SR proteins from HepG2 cell line. TPM values were obtained from the control group which were treated with control shRNA. Color of the boxes correspond to shuttling activity of the SR protein. SRSF5 has the highest transcript expression in HepG2 among all SR genes.

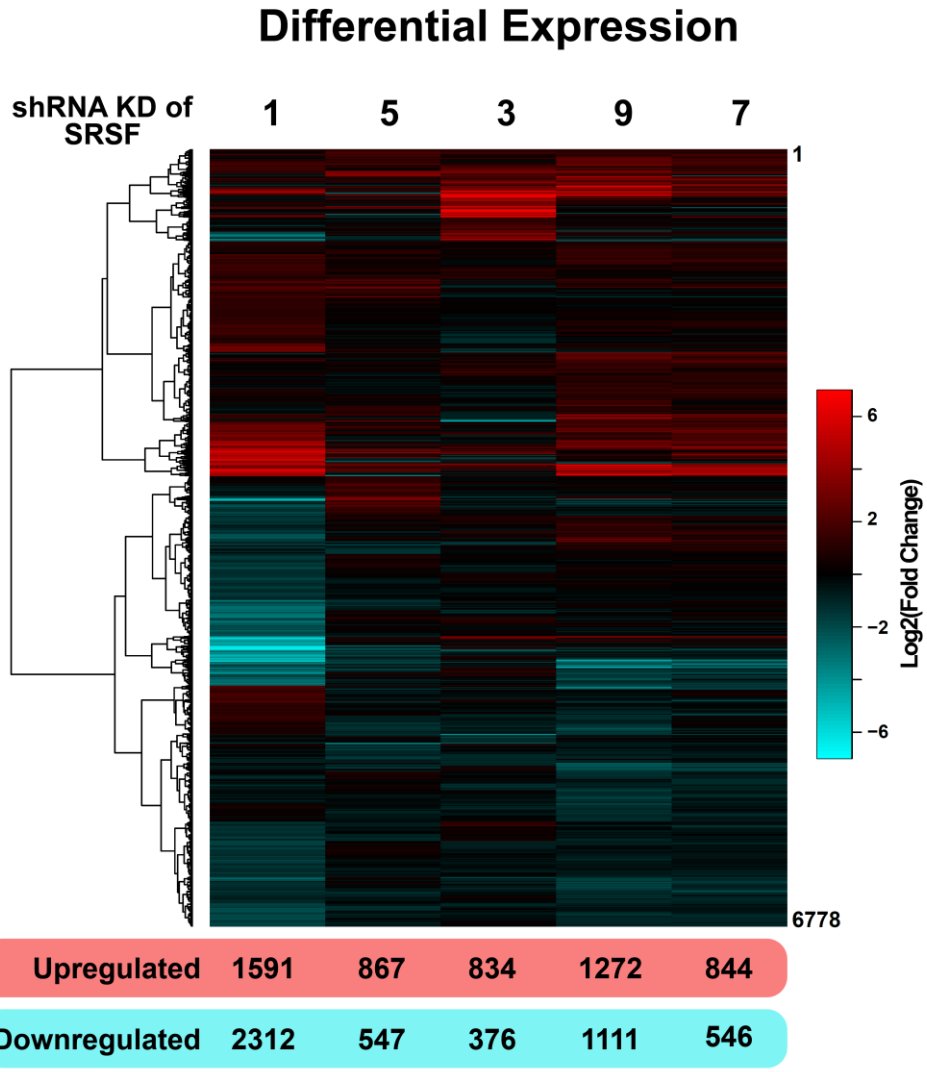


Figure 7.2: ***Gene Expression changes in SR Protein Knockdown in HepG2***

Thousands of genes show changes in gene expression upon knockdown of SR proteins in HepG2. SRSF1 shows the greatest number of gene expression changes. There are a greater number of genes being downregulated upon SRSF1 knockdown.

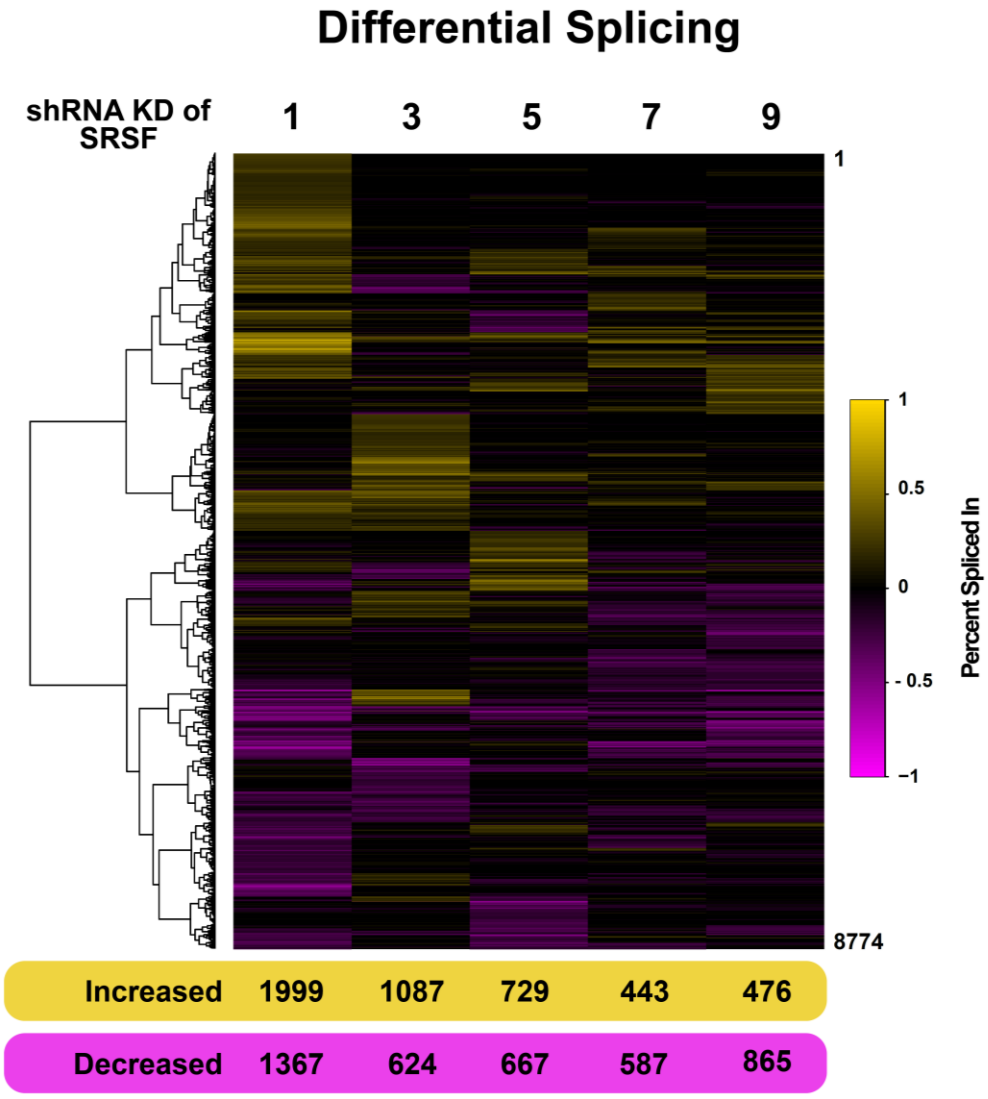


Figure 7.3: ***Splicing Changes in SR protein knockdown in HepG2***

Thousands of events undergo differential splicing upon knockdown of SR proteins in HepG2. SRSF1 shows the greatest number of events changing in splicing.

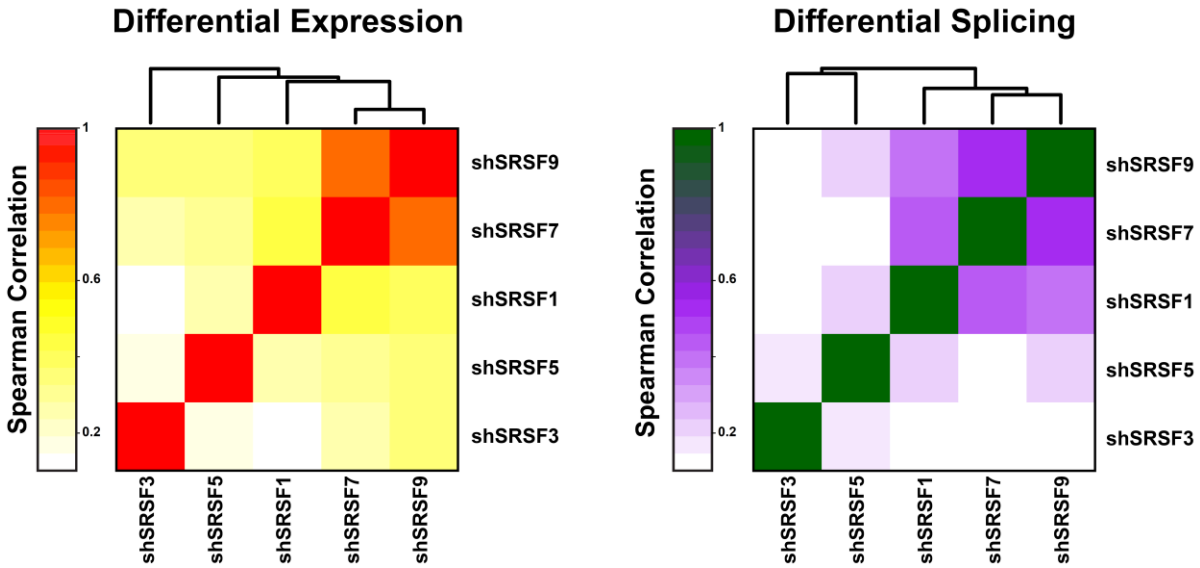


Figure 7.4: ***SRSF1 is Most Similar to SRSF7 and SRSF9 in Regard to Expression and Splicing***

Heatmap of the pairwise spearman correlation values for differential expression and splicing in shRNA knockdown of SR proteins in HepG2. SRSF1 appears to be more correlated to SRSF7 and SRSF9.

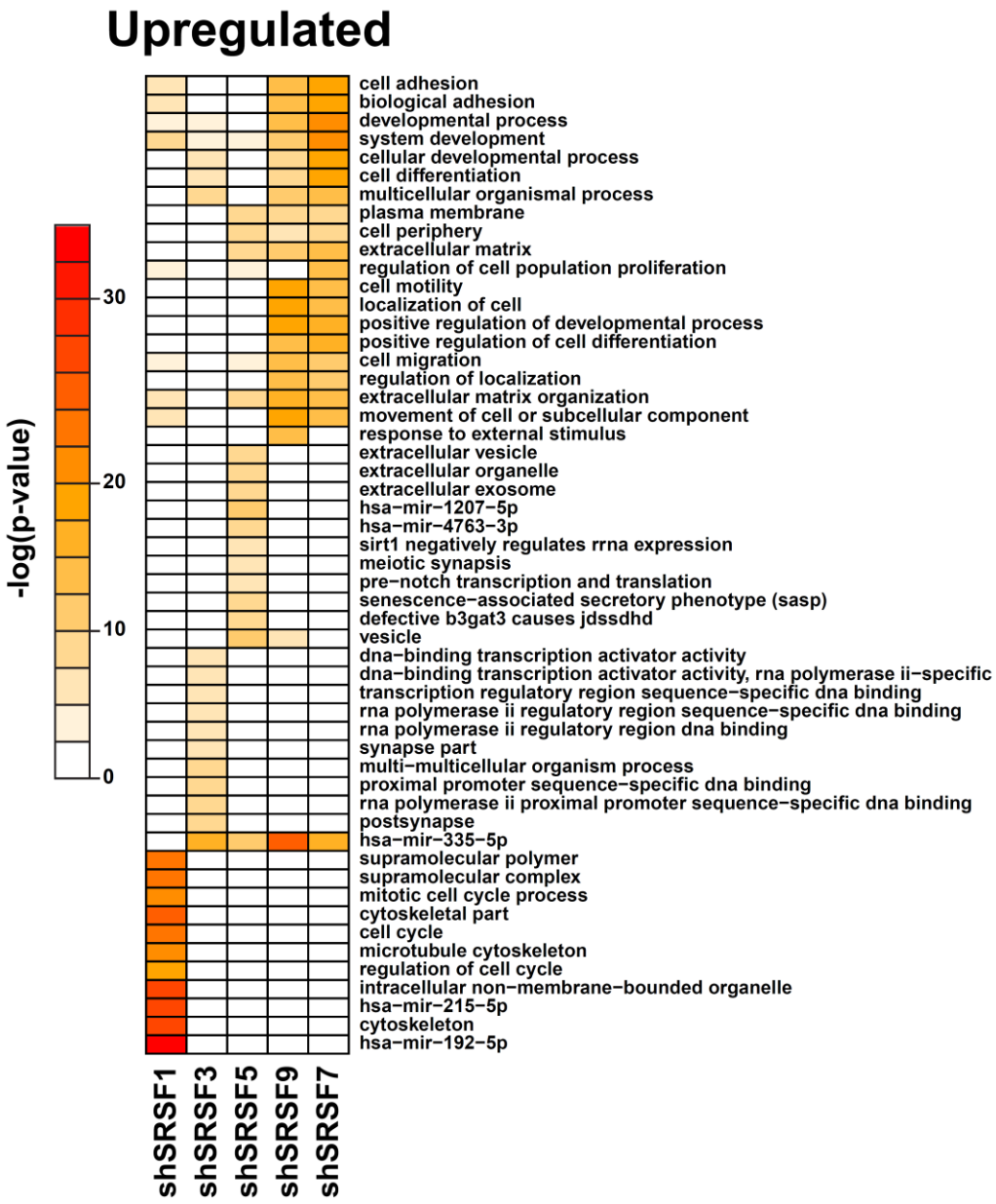


Figure 7.5: ***Knockdown of SRSF1 Triggers Upregulation of Genes Involved in Cell Cycle Processes***

Heatmap of Gene Ontology enrichment terms based on -log(p-value) for genes upregulated in the associated SR gene knockdown by shRNA. SRSF1 knockdown triggers upregulation of cell cycle genes.

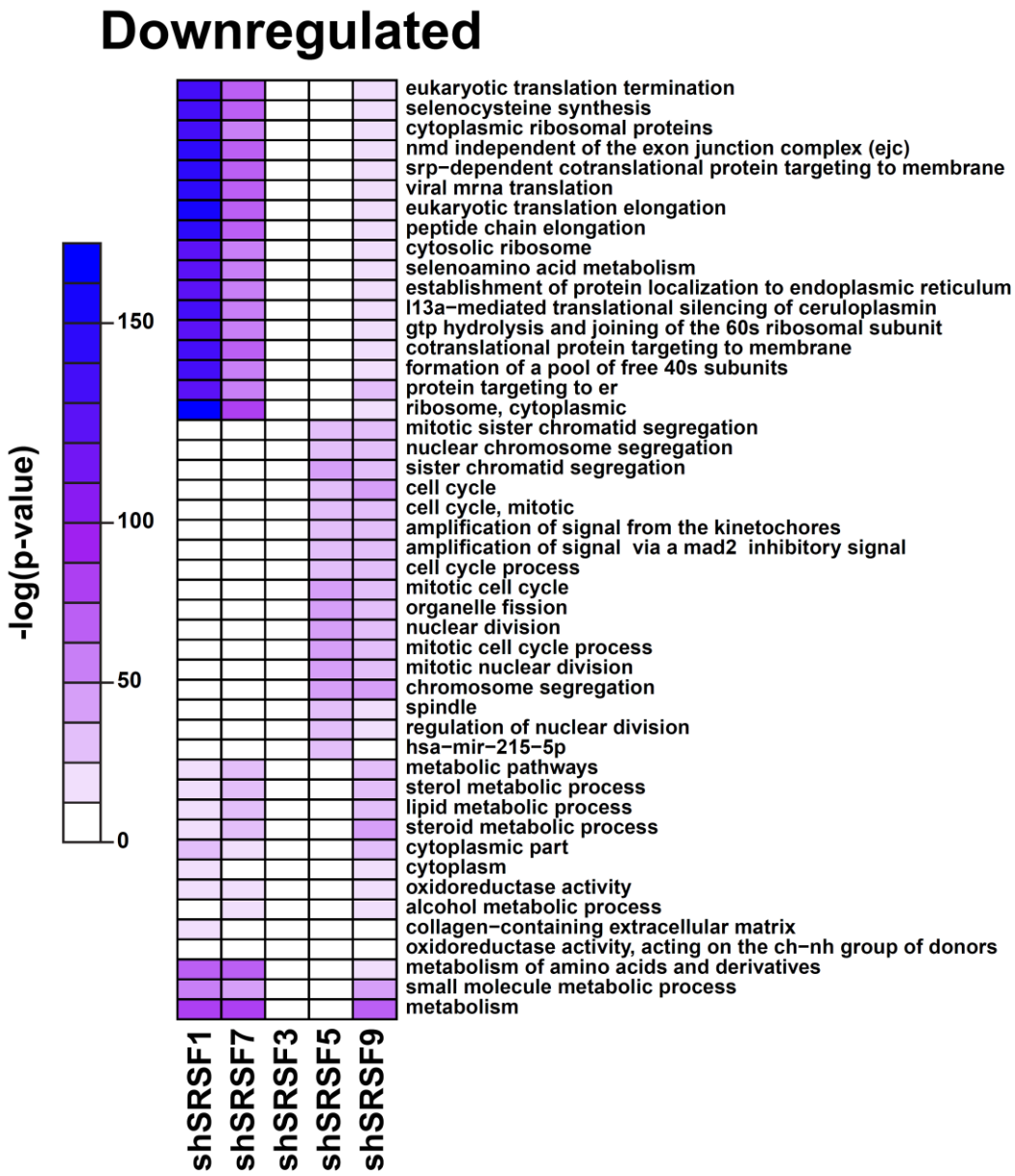


Figure 7.6: ***Knockdown of SRSF1 in HepG2 Results in Downregulation of Ribosomal and Translation Regulation Genes***

Heatmap of Gene Ontology enrichment terms based on -log(p-value) for genes downregulated in the associated SR gene knockdown by shRNA. SRSF1 knockdown exhibits strong enrichment for ribosomal genes and genes involved in various aspects of translation.

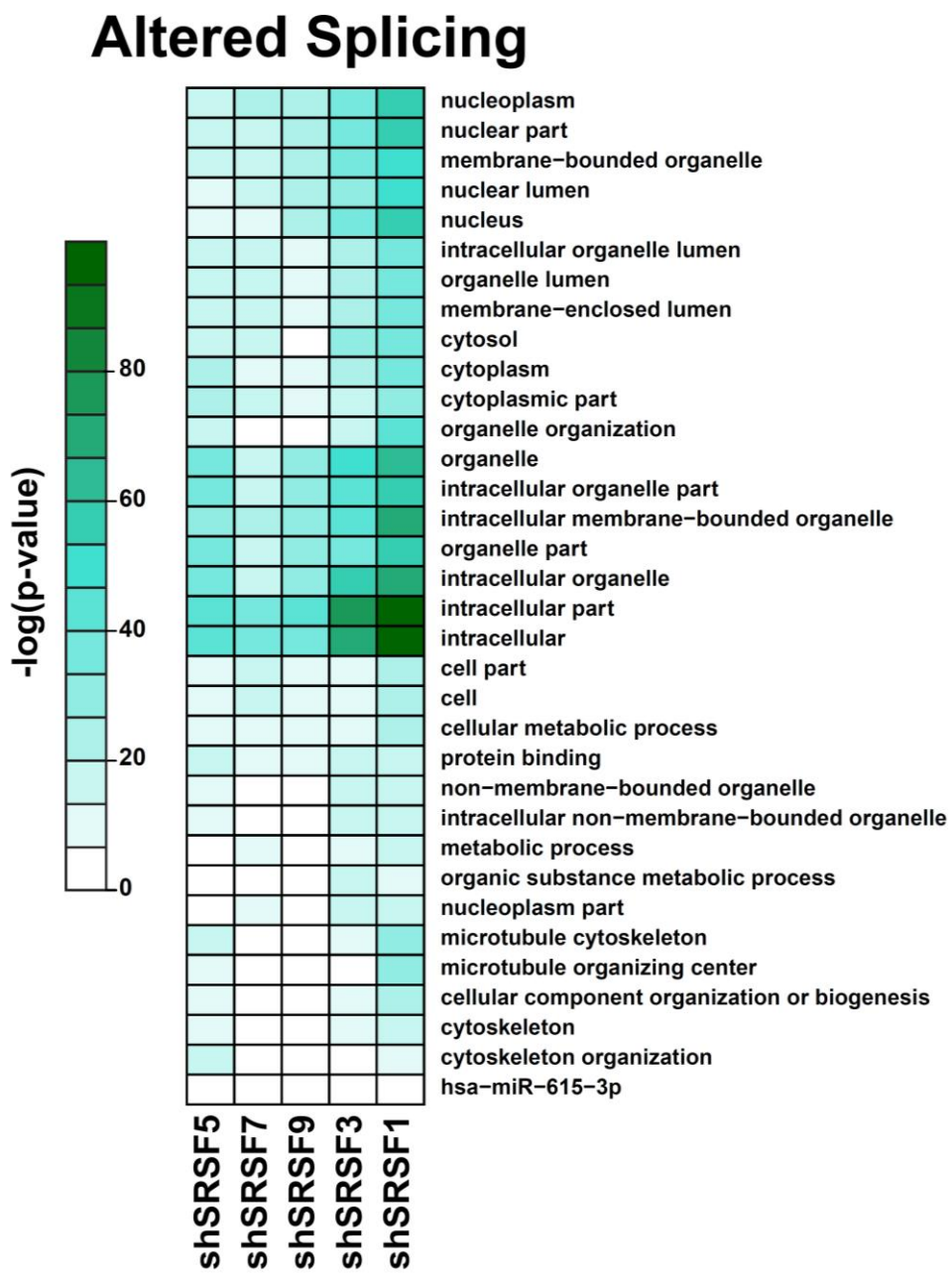


Figure 7.7: ***Gene Ontology Analysis for Differential Splicing in SR Gene Knockdown***

Heatmap of Gene Ontology enrichment terms based on -log(p-value) for genes with differential splicing in the associated SR gene knockdown by shRNA. GO term enrichment was not strong for any biological process in any SR gene knockdown dataset.

Knockdown in HepG2	Analysis of Differential Expression (DESeq2)	Splicing (rMATS)	Intron Retention (IR Finder)
SRSF1	1591 2312	1999 1367	547 51
SRSF3	834 376	1087 624	17 6
SRSF5	867 547	729 667	0 2
SRSF7	844 546	443 587	329 5
SRSF9	1272 1111	476 865	739 7

Figure 7.8: ***Summary Table of Expression, Splicing, and Intron Retention Analysis in SR Gene Knockdown RNA-seq in HepG2 from the ENCODE Project***

Table showing the number of genes or events with differential expression or splicing in the indicated SR gene knockdown. Numbers in red signify increased expression or exon inclusion while blue signify decreased expression or exon skipping events. Of all SR gene knockdowns, SRSF1 has the largest affect on the transcriptome.

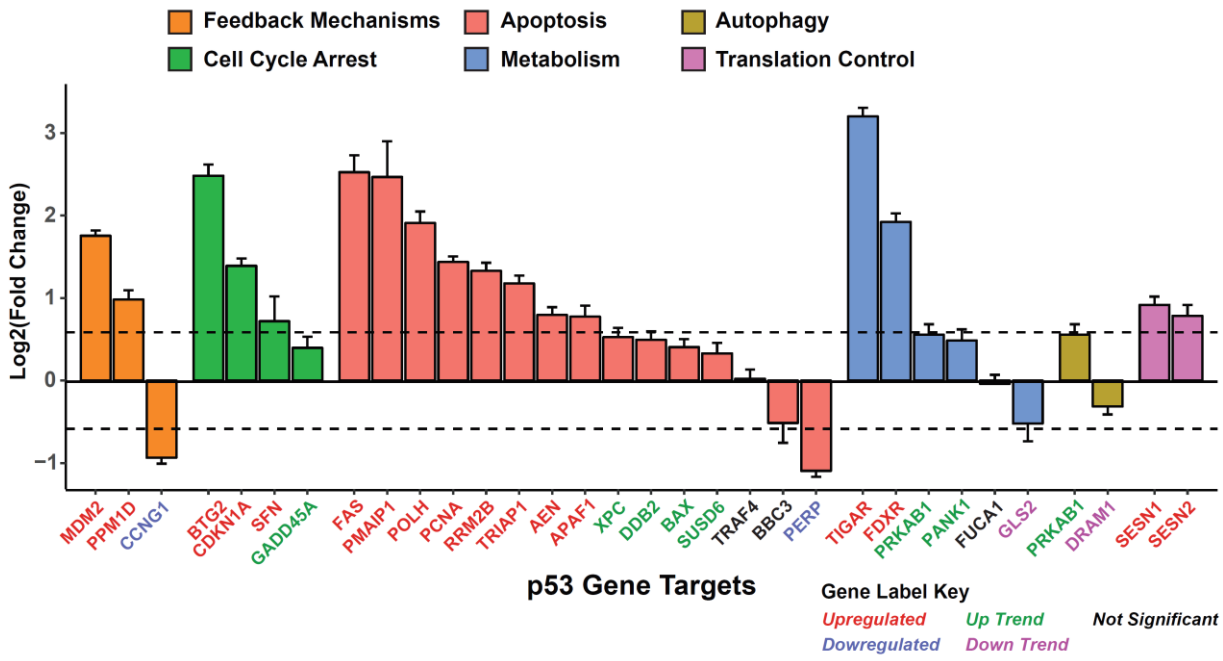


Figure 7.9: ***Knockdown of SRSF1 in HepG2 Triggers p53 Response***

Bar plot showing changes in expression estimated from RNA-seq of various p53 target genes upon knockdown of SRSF1 in HepG2. Multiple p53 target genes are strongly induced with knockdown of SRSF1.

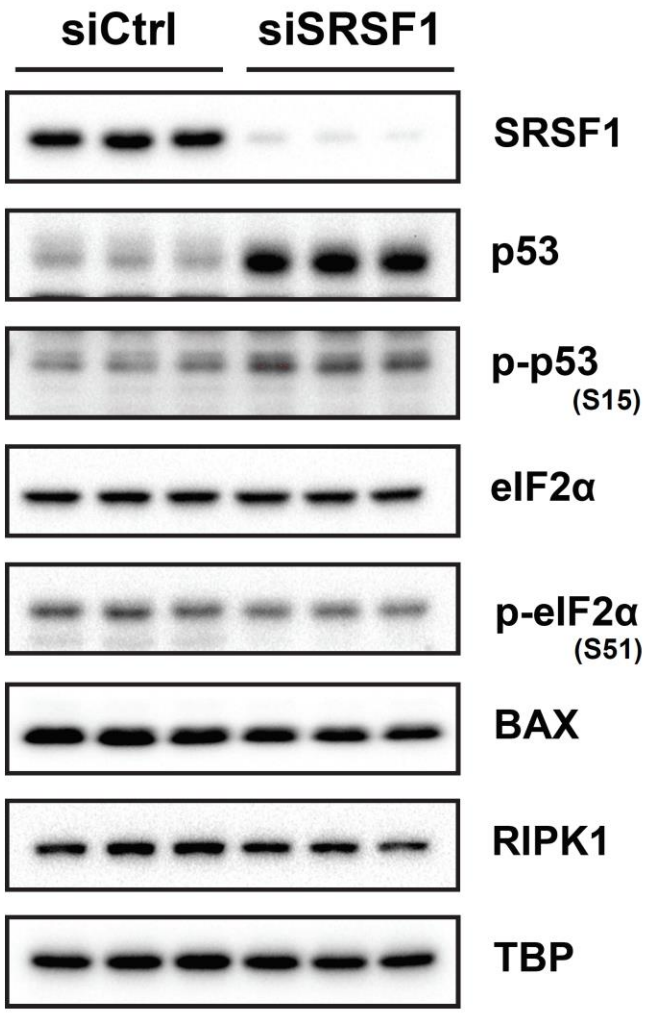


Figure 7.10: ***Knockdown of SRSF1 in HepG2 Triggers Induction of p53 Levels***

Western blot analysis of indicated factors after 72 hours of specified siRNA treatment in HepG2 cells. Blots depict representative samples for each group (n = 4 – 5 per group). Efficient knockdown of SRFS1 is achieved. Levels of p53 are striking induced in the siSRSF1 treated HepG2 cells. Induction of BAX or RIPk1 are not induced signifying cell death has not been triggered yet.

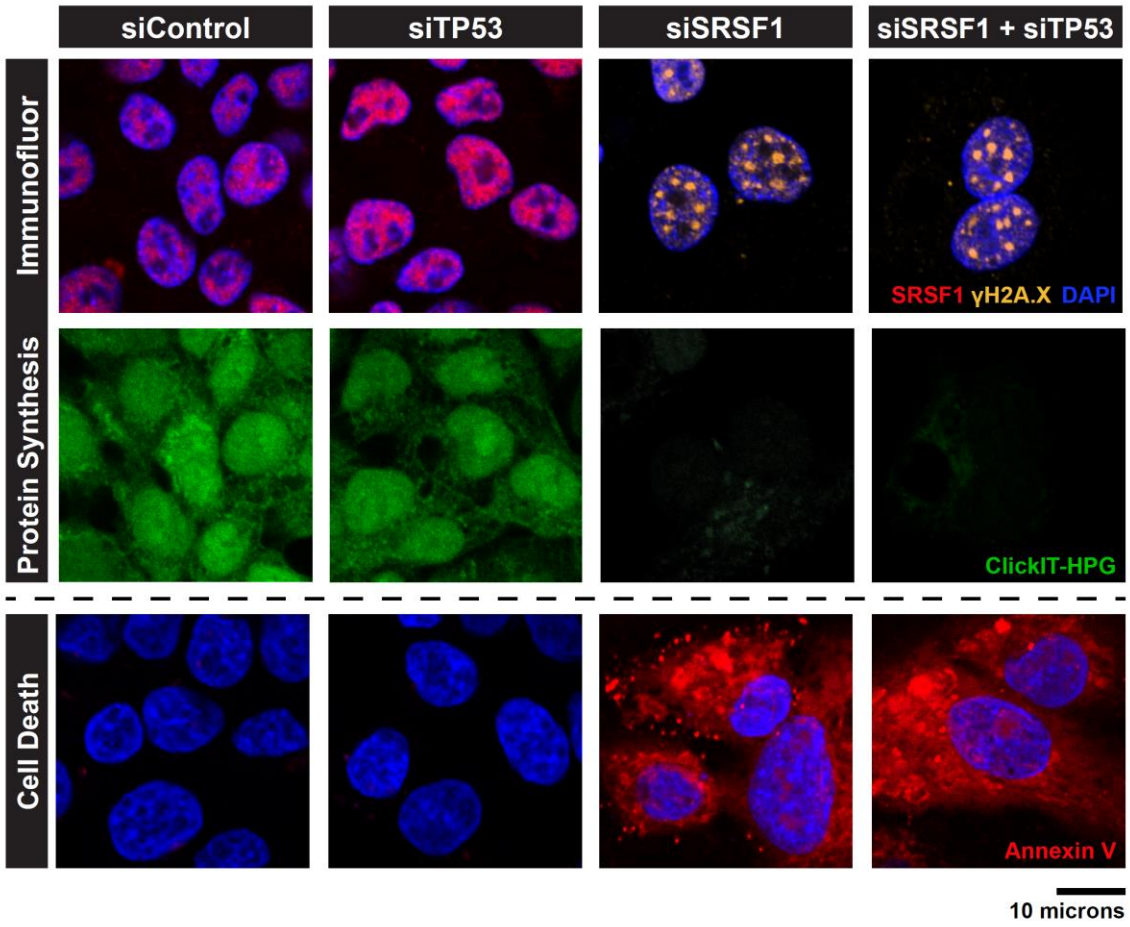


Figure 7.11: ***Loss of SRSF1 Activity in HepG2 Induces DNA Damage, Decreased Protein Synthesis, and Cell Death and is Independent of p53***

HepG2 cells were cultured and treated with the indicated siRNA for 72 hours before performing subsequent assays. The first row of images shows representative images (n = 5 fields by 3 replicates per group) of immunofluorescent staining for SRSF1 (red), γH2A.X (yellow), and nuclear (blue). The second row corresponds to nascent protein synthesis signal (green) of the same field using a fluorescent based ClickIT-HPG incorporation assay. Knockout of SRSF1 results in DNA damage and inhibition of protein synthesis. This response is independent of p53 activation. The last row shows representative images (n = 5 fields by 3 replicates per group) of Annexin V staining of HepG2 cells 96 hours post knockdown. Loss of SRSF1 results in eventual cell death.

Chapter 8: Summary

The primary motivation of this project was to understand the role of SRSF in liver physiology. Investigation of its role was started by generating and studying hepatocyte-specific SRSF1 knockout mice models. It was well established that whole-body knockouts of SRSF1 results in embryonic lethality in multiple model organisms. Therefore, it was a shocking surprise that the SRSF1 HKO mice were able to maintain viable livers at all. Although the livers displayed significant inflammation and damage, SRSF1 HKO mice must have some degree of liver function to be able to survive for as long as they do. Our assumption that SRSF1 was efficiently knocked out of the hepatocytes in this model was incorrect. It was not realized until later that hepatocytes were able to subvert the knockout of SRSF1 by repressing the expression of the Cre transgene. Due to this incorrect understanding of the model, the initial working hypothesis explaining the observed phenotype was that loss of SRSF1 was causing changes in the transcriptome resulting in the observed inflammation, steatosis, and fibrosis **(Figure 8.1)**.

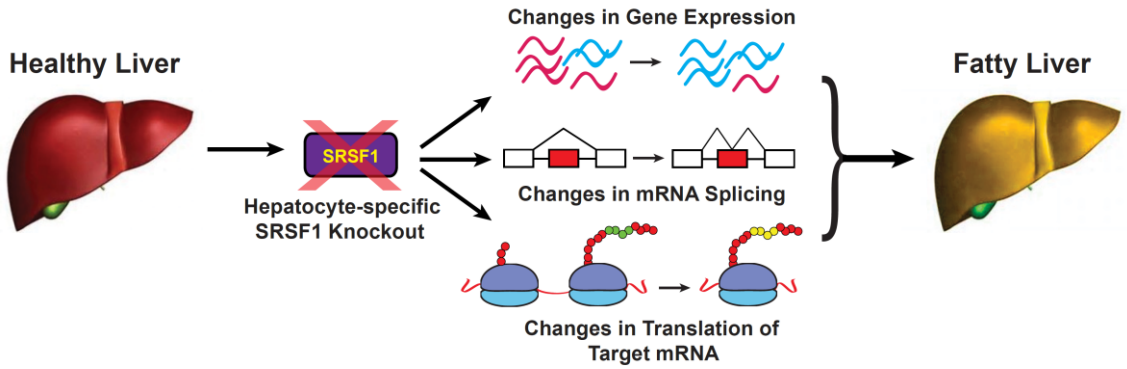


Figure 8.1: ***Initial Working Hypothesis of Hepatocyte-specific SRSF1 Knockout Model***

As a result of this initial working hypothesis, much of the efforts in the beginning were focused on studying the irregularities in the transcriptome. This resulted in performing multiple RNA-seq and eCLIP-seq experiments. However, these studies did not provide viable explanations to the phenotype that was observed in the SRSF1 HKO model. This eventually led to the development of the second SRSF1 knockout model which is referred to as the Acute SRSF1 HKO model. The motivation behind developing this model was that an acute viral-mediated knockout of SRSF1 in an adult animal will provide additional temporal control over the deletion of SRSF1. We reasoned that a major limitation of the SRSF1 HKO model was that pathological changes were occurring too quickly. Therefore, insight into the pathological mechanism was being obscured due to the overwhelming secondary effects which was masking the primary effectors. To overcome this issue, the acSRSF1 HKO will allow for isolation of hepatocytes before any secondary changes could take place. While this model provided the ability to study SRSF1 knockout prior to secondary changes, like the SRSF1 HKO, transcriptome studies on this model failed to offer any additional insights into the liver pathology of HKOs.

Progress into the mechanistic understanding began to change when the immunofluorescent staining assay for SRSF1 in liver tissue was developed. This assay allowed for identifying the cells expressing SRSF1 within the intact liver tissue. Shockingly, when this assay was performed on the SRSF1 HKO tissue it was clear that the liver was repopulating the tissue with SRSF1-expressing hepatocytes. It was eventually realized that this was due to proliferation of hepatocytes where Cre expression is suppressed thus allowing SRSF1 to stay intact. This phenomenon is also not novel and has been observed previously with the generation of the hepatocyte-specific Dicer1

knockout mice (Sekine et al. 2009). Like SRSF1 HKO, efficient deletion of Dicer1 was achieved in young mice liver followed by repopulation with Dicer1-expressing hepatocytes. Insight into the mechanism of how the livers can bypass the knockout was shown by Duncan, et al 2012. They show that because of selection pressure, expansion of knockout resistant aneuploid hepatocytes occurs. In the case of SRSF1 HKO, it is likely that the Cre transgene locus is being lost leading to the generation of SRSF1 knockout resistant hepatocytes. With this finding it was understood that SRSF1 was vital the functioning of the cell and loss of its activity was resulting in immediate death of the cell.

At this point, extensive review of the past literature on SRSF1 was performed to identify any studies which investigated its knockout in a system. This literature review led to a 2005 paper by Li and Manley where they developed a tet responsive SRSF1 knockdown system in DT40 cells, a chicken B cell line. Surprisingly, they discovered that after SRSF1 depletion, the cultures eventually began to grow cells that were resistant to tet-mediated knockdown. They ultimately discovered that loss of SRSF1 resulted in the accumulation of R-loops leading to subsequent DNA damage. It was quickly realized that a similar phenomenon must be occurring in the SRSF1 HKO models. Indeed, probing of SRSF1 HKO tissue sections with γH2A.X, a well-established DNA damage marker, showed robust signal within the nuclei. Furthermore, dot blot assay for the presence of R-loops revealed a remarkable accumulation within the SRSF1 HKO. Therefore, loss of SRSF1 is leading to R-loop accumulation with subsequent DNA damage.

An additional observation made during phenotypic characterization of the SRSF1 HKO model was that these mice exhibited significant impairment of global translation. Proteomics analysis revealed striking downregulation of ribosomal proteins and factors

involved in translation. Moreover, polysome profiling showed absence of polysomes while a puromycin incorporation assay revealed diminished protein synthesis upon knockout of SRSF1. A possible explanation for this finding is that the extensive DNA damage in SRSF1 HKO is resulting in activation of p53 which triggers the subsequent shutdown of global translation. However, western blot analysis of SRSF1 HKO does not indicate activation of p53. Interestingly, when knockdown is performed in HepG2 cells, activation of p53 is observed along with DNA damage and inhibition of protein synthesis. A double knockdown experiment was conducted where both SRSF1 and p53 were simultaneously depleted to check if it could rescue the translation shutdown response. The findings from this experiment indicated that the protein synthesis inhibition following SRSF1 knockdown is independent of p53.

To reiterate, this project's main goal was to understand the role of SRSF1 in liver physiology. Investigation of the SRSF1 HKO model led to the understanding that SRSF1 is vital for the overall physiology of all eukaryotic cells. This is due to its crucial role in protecting genomic DNA by preventing the formation of deleterious R-loops between nascent RNAs and template DNA during transcription. In the setting of SRSF1 knockout, R-loop accumulation induces DSBs which leads to inhibition of global translation and subsequent death of the cell. In the context of the liver tissue, widespread hepatocyte death results in hepatic failure, which encompasses decreased accumulation of triglycerides within the tissue, failed formation of lipoprotein particles, and impaired levels of serum glucose and ketone bodies **(Figure 8.2)**.

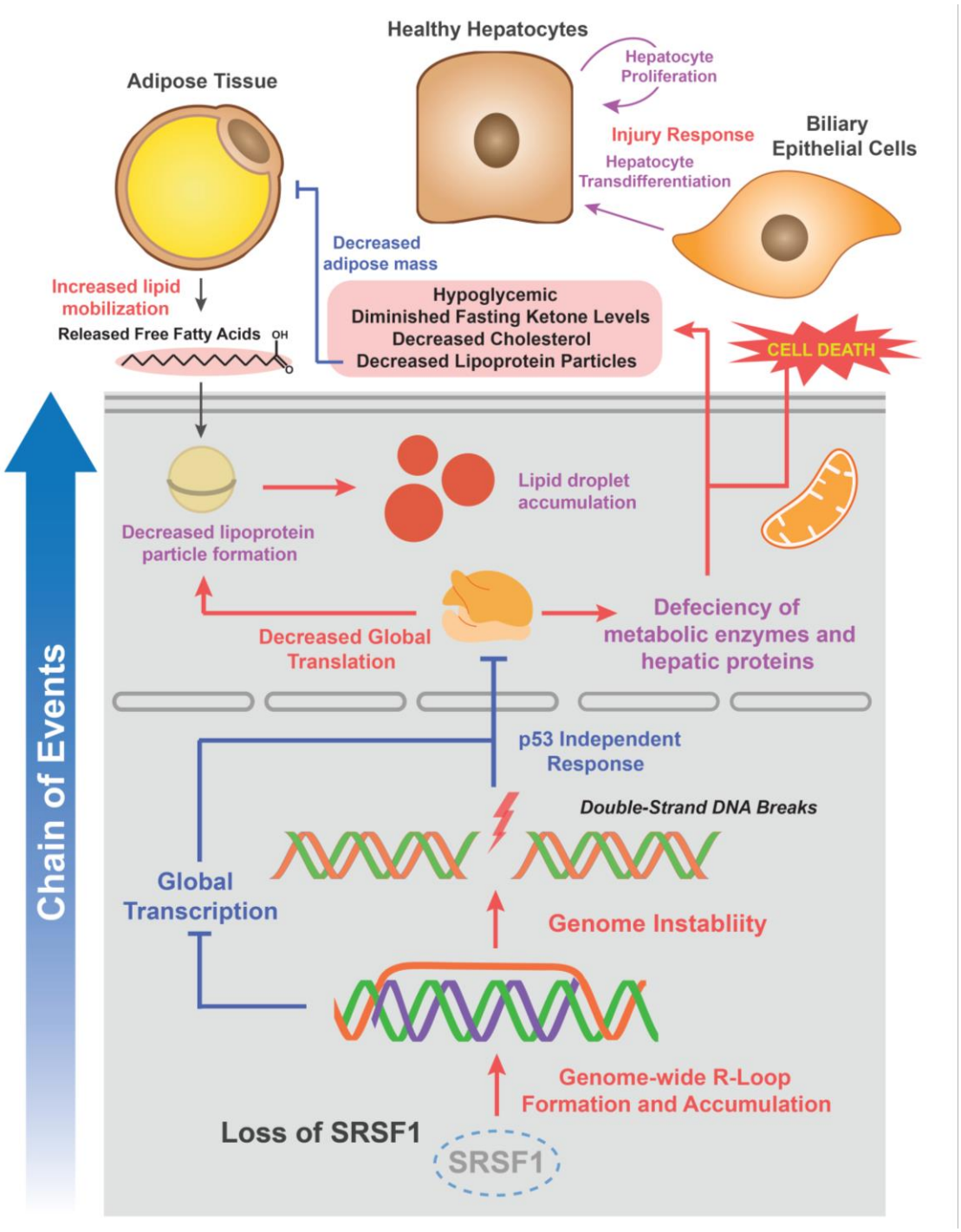


Figure 8.2: ***Proposed Mechanism of Hepatocyte Response Upon SRSF1 Knockout***

Loss of SRSF1 results in increased accumulation of unresolved R-Loops resulting in DNA damage through double stranded breaks. The genome instability triggers a response resulting in global inhibition of translation. Without SRSF1 the cell is not able to resolve the irregularities in gene expression and results in subsequent failure of the cell.

Chapter 9: Supplementary Information

9.1 Supplemental Protocols

9.1.1 Genotyping of SRSF1 HKO Mice Model Using Tail Clippings

Mouse genotyping was performed on genomic DNA isolated from tail clippings. Mice tails were lysed using 100 uL of DirectPCR tail lysis buffer (Viagen Biotech) supplemented with Proteinase K (20 µg) and incubated overnight at 55 °C. The following day, lysed tails were incubated at 100 °C for 20 minutes to heat inactivate Proteinase K. The mixture was then spun down to pellet insoluble debris and 1 µL of the supernatant containing genomic DNA was used to setup PCR reactions. Primers, reaction setup and cycling protocol are detailed in the following tables.

Table 9.1: Genotyping Primer Sequence

Genotyping Primer Set	Forward Primer (5' → 3')	Reverse Primer (5' → 3')
SRSF1 Flox	GGGACTAATGTGGGAAGAATG	AACCTAAACTATTGCTCCCATCTG
Albumin-CRE	GAAACCTGATGGACATGTTCAGG	AGTGCGTTCGAACGCTAGAGCCTGT

Table 9.2: SRSF1 Genotyping PCR Reaction Component Setup

SRSF1 Flox Genotyping Reaction Setup	
Components	Volume/Reaction (µL)
10X PCR Buffer	2
50 mM $MgCl_2$	0.6
10mM dNTPs	0.4
Platinum Taq	0.08
Forward Primer, 10 µM	0.5
Reverse Primer, 10 µM	0.5
Water	14.92
DNA	1
Total	20

Table 9.3: AlbCre Genotyping PCR Reaction Component Setup

AlbCre Genotyping Reaction Setup	
Components	Volume/Reaction (µL)
5X Taq Mix	5
Forward Primer, 10 µM	0.5
Reverse Primer, 10 µM	0.5
Water	18
DNA	1
Total	25

Table 9.4: SRSF1 Flox Genotyping Reaction Thermocycling Conditions

SRSF1 Flox Genotyping PCR Thermocycling Conditions		
Step	Temperature	Time
Initial Denaturation	94°C	2 minutes
10 Cycles	94°C 65°C 68°C	20 seconds 15 seconds 10 seconds
28 Cycles	94°C 50°C 72°C	15 seconds 15 seconds 10 seconds
Final Extension	72°C	2 minutes
Hold	4-10°C	

Table 9.5: AlbCre Genotyping Reaction Thermocycling Conditions

AlbCre Genotyping PCR Thermocycling Conditions		
Step	Temperature	Time
Initial Denaturation	94°C	3 minutes
30 Cycles	94°C 52°C 72°C	30 seconds 40 seconds 30 seconds
Final Extension	72°C	3 minutes
Hold	4°C	

Genotyping reactions were then separated on a 5% PAGE gel, stained with ethidium bromide, and visualized using the ChemiDoc XRS+. The following gel images depicts a representative genotyping gel and its interpretation.

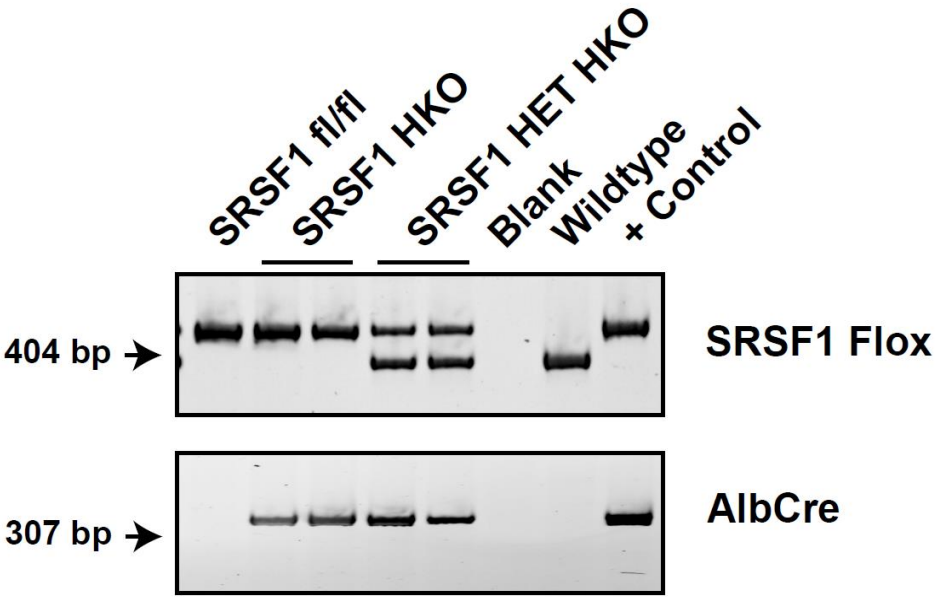


Figure 9.1: ***SRSF1 HKO Mice Model Genotyping Gel***

SRSF1 HKO mice model genotyping reactions separated on 5% PAGE gel.

9.1.2 Purification of Hepatocytes from Mice Liver

Hepatocytes were isolated and purified using the Two-Step Collagenase Perfusion Technique (W. C. Li, Ralphs, and Tosh 2010). Mice were first anesthesized with isoflurane and then secured to a surgery pad with the ventral side up. A “U” shaped incision was made on the abdomen to expose the liver. The liver was then perfused through the portal vein with 50 ml of wash buffer containing 1X Hanks Balanced Salt Solution (HBSS) and 1 mM EDTA (pH 8.0), without calcium and magnesium salts. Following this, the livers were perfused using 50 ml of digestion buffer containing 1X HBSS, 0.5 mM $CaCl_2$, 40 µg/mL soybean trypsin inhibitor, and 60 U/mL of Collagenase

Type I from Worthington. The perfused liver was carefully excised out from the abdomen and transferred into a petri dish containing 1X HBSS. Using cell scrapers, the tissue was carefully massaged to release the cells from the capsule. The crude cell prep was then filtered through a 40 µm mesh filter and the resulting single cell suspension was centrifuged at 50 g for 5 minutes at 4 °C. The supernatant containing non-parenchymal cells (NPCs) and dead hepatocytes was discarded while the pellet containing live hepatocytes were resuspended in fresh 1X HBSS. The centrifugation wash was repeated two additional times before being aliquoted into 1.5 mL microcentrifuge tubes, flash frozen in liquid nitrogen and stored in -80 °C until further use.

To assess whether efficient separation of hepatocytes and NPCs was achieved, RT-PCR followed by gel electrophoresis assay was performed as described in *Section 9.1.6* for various hepatocyte and NPC markers. The following figure demonstrates that efficient separation of hepatocyte and NPC cell populations was achieved using this protocol **(Figure 9.2)**. Specifically, multiple NPC markers are absent in the hepatocyte fraction and vice versa.

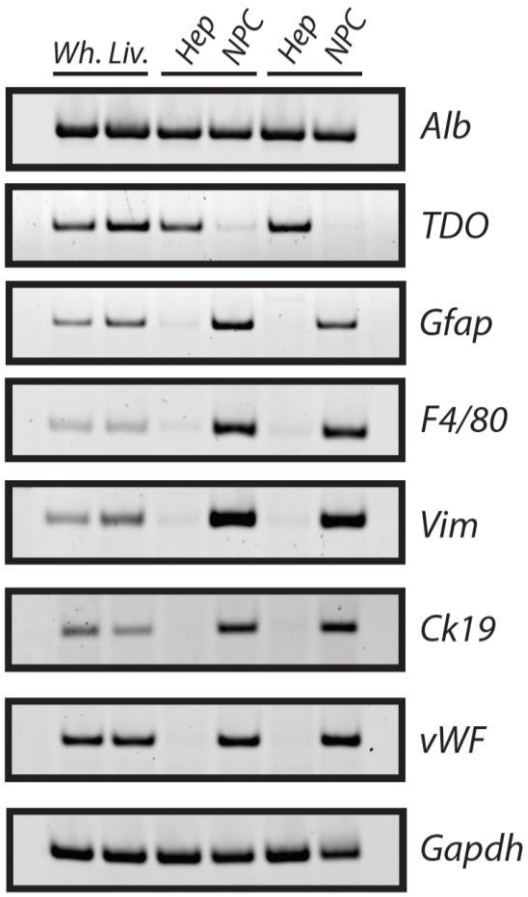


Figure 9.2: ***Assessment of Hepatocyte Isolation from Mice Whole Liver Tissue***

Separation of hepatocytes and NPCs was achieved with significant enrichment for the associated population. Hepatocyte makers: Alb, TDO; NPC markers: Gfap, F4/80, Vim, Ck19, vWF; Control marker: Gapdh.

9.1.3 Protein Isolation from Tissue and Western Blotting Analysis

Total proteins were isolated from ~50 mg of snap frozen liver tissue or purified cell pellet by homogenizing in 400 µL of cold homogenization buffer containing 10 mM HEPES (pH 7.5), 0.32M Sucrose, 5 mM EDTA, 1% SDS, 5 µM MG132 and Pierce Protease Inhibitors (1 tablet per 10 mL of buffer, Catalog # A32953). Samples were sonicated in a water bath to shear DNA and clarified by centrifugation. Protein concentration was

determined using the Pierce™ BCA Protein Assay Kit (Thermo Scientific). Protein lysates were diluted to 5 mg/mL and then boiled in 1X Laemmli buffer at 100 °C for 10 minutes. Once boiled, samples were cooled to room temperature and ~50 µg of proteins were resolved on a 10% SDS-PAGE gel and transferred using a wet transfer setup onto a PVDF membrane with 0.45 µm pore size (Immobilon, Millipore). Membranes were blocked using Tris-buffered saline containing 5% nonfat dry milk and 0.1% Tween 20 (TBST). After blocking, membranes were incubated with primary antibody overnight at 4°C. The membranes were then washed with TBST to remove any unbound primary antibody followed by incubation with an appropriate horseradish peroxidase-conjugated secondary antibody for two hours. Membranes went through a final TBST wash and then visualized on the ChemiDoc XRS+ using the Clarity Western ECL kit (BioRad). Refer to *Section 9.1.9* for details on antibody source and dilutions.

9.1.4 Histological, Immuno-histochemistry and fluorescence Staining

For paraffin embedding, liver tissues were harvested and immediately fixed in 10% neutral-buffered formalin overnight at 4 °C. Fixed tissues were then processed in a series of solvents starting from ethanol solutions, to xylenes, and finally embedded into paraffin wax. Paraffin embedded tissues were cut into 5 µm thick sections using a microtome. The sections were placed onto charged microscope slides and allowed to air dry overnight before performing histological staining. For frozen tissue embedding, a piece of fresh and unfixed tissue was submerged into a plastic mold filled with OCT medium (Tissue-Tek, Sakura) and frozen using liquid nitrogen. Using a cryostat, sections were cut into 10 µm thick sections and transferred onto a room temperature microscope slide. The slides with frozen tissue sections were stored at -80 °C prior to staining.

- Hematoxylin and Eosin Staining

Slides with paraffin embedded tissue sections were deparaffinized in xylene and then rehydrated by incubating them in a decreasing gradient of ethanol solutions with a final incubation in distilled water. Once rehydrated, the sections were stained with Modified Harris Hematoxylin Solution (7211 Richard-Allan Scientific) for 2 minutes. It was then briefly rinsed in water and then stained with Eosin-Y Solution (7111 Richard-Allan Scientific) for 1 minute. The stained sections were then dehydrated by incubating the slides in an increasing gradient of ethanol solutions with a final incubation in xylene. Once dehydrated the slides were mounted with Permount (Fisher Scientific) mounting media and coverslip.

- Sirius Red Staining

Slides with paraffin embedded tissue sections were deparaffinized and rehydrated as explained in the *Hematoxylin and Eosin Staining* protocol. Slides were then incubated in Picro-sirius red solution (0.1% w/v Direct Red 80 in saturated aqueous solution of picric acid) for 1 hour. They were then briefly rinsed in two changes of 0.5% acetic acid solution. The slides were then dehydrated and mounted in the same manner as described in the the *Hematoxylin and Eosin Staining* protocol.

- Oil Red O Staining

Frozen tissue sections were taken out from -80 °C and allowed to air dry for about 5 minutes. They were then fixed in 10% Neutral-buffered formalin solution for 10 minutes at room temperature. Slides were then briefly rinsed in tap water and then 60% isopropanol solution. Slides were then stained with freshly prepared Oil Red O solution

(0.3% w/v Oil Red O in 60% isopropanol) for 15 minutes. Once stained, the slides were rinsed in 60% isopropanol solution and gently counterstained with Modified Harris Hematoxylin. Slides were rinsed with distilled water and mounted using CC mount (C9368 Sigma) and coverslip. Slides were allowed to dry before imaging.

- Immunohistochemistry (IHC) and Immunofluorescent (IF) Staining

Slides with paraffin embedded tissue sections were deparaffinized and rehydrated as explained in the *Hematoxylin and Eosin Staining* protocol. Slides were then antigen retrieved in buffer containing 10 mM Tris, 1 mM EDTA, 0.05% Tween 20, pH 8.0 for 5 minutes at 120 °C using a slow cooker. The sections were then incubated in wash buffer containing 1X TBS and 0.025% Triton X-100 and then blocked using 1X TBS, 10% Normal Goat Serum (NGS) and 1% BSA for 2 hours at room temperature (RT). Primary antibodies were applied to the sections at standardized concentrations and incubated overnight at 4 °C. Following this, the sections were incubated in wash buffer, and secondary HRP-conjugated or fluorescent antibodies were applied for 1 hour at RT for IHC or IF, respectively.

IHC - The sections were washed in 1X TBS and then developed using a DAB Peroxidase Substrate kit (Vector Laboratories) for approximately 5 minutes. Sections were counterstained with hematoxylin, dehydrated into xylenes, and then mounted using Permount.

IF – The sections were washed in 1X TBS and then stained for the nucleus using ToPro3 for 15 minutes at RT. The sections were then cover slipped using CC aqueous mounting

media. All sections were imaged on a Zeiss LSM 710 microscope at IGB core facility, UIUC. All antibodies used, and respective dilutions are listed in Section 9.1.9.

9.1.5 Total RNA Isolation and Real Time Quantitative RT-PCR (qRT-PCR) Analysis

Total RNA was isolated from about 50 mg of snap frozen liver tissue or snap frozen cells from 6-well culture plates with TRIzol (Invitrogen) using the protocol described in the manual. Quality of the RNA was assessed by running about 1 µg of RNA on a bleach gel to evaluate the 28S and 18S bands (Aranda, LaJoie, and Jorcyk 2012). Approximately 5 µg of RNA was then reverse-transcribed into cDNA using the Maxima Reverse Transcriptase (Thermo Fisher Scientific) following manufactures protocol. Relative gene expression analysis was performed with ~50 ng of cDNA per reaction using a SYBR® Green™ based assay for Real-Time Quantitative Reverse Transcription PCR (qRT-PCR). The following table details the cycling conditions for the reaction.

Primer sets used to assay the various genes in the study are listed in *Primer Sequence Table* in Section 9.1.8. Using *36B4* as a loading control, relative gene expression compared to the control group was calculated by the double Ct method (Livak and Schmittgen 2001).

9.1.6 Reverse Transcription PCR followed by Gel Electrophoresis Analysis

Total RNA and subsequent preparation of cDNA was performed as specified in *Section 9.1.5*. The cDNA was diluted to 25 ng μl^{-1} with nuclease free water from which 1 µL would be used for PCR reactions. PCR reactions were setup using Taq DNA Polymerase (NEB #M0237) along with forward and reverse primers to a final concentration of 0.2 µM. Cycling conditions used for the PCR reaction are as follows; Initial denaturation at 95 °C for 30 seconds, 30 cycles consisting of denaturation at 95 °C

for 20 seconds, annealing at 60 °C for 30 seconds, and extension at 68 °C for 30 seconds, and a final extension at 68 °C for 5 minutes. For visualization, 10 µL of reaction mixture was mixed with loading dye and then resolved on a 5% PAGE gel, stained in an ethidium bromide bath, and then imaged using ChemiDoc XRS+.

9.1.7 TUNEL Staining on Tissue Sections for Apoptosis

TUNEL staining was performed on paraffin tissue sections that were deparaffinized and rehydrated as specified in the *Hematoxylin and Eosin Staining* protocol in Section 9.1.4. Staining was performed with the *In-Situ Cell Death Detection Kit, Fluorescein* (Sigma Cat# 11684795910) using manufacturer protocol. Briefly, rehydrated tissue sections were permeabilized and then labelled using the TUNEL reaction mixture from the kit. The sections were incubated with the reaction mixture for 1 hour at 37 °C in a humidified chamber. Slides were then rinsed with PBS three times, counterstained for nucleus, and mounted with coverslip. For positive control, the section was treated with 20 U of DNAse (NEB) mix for 15 minutes. The slides were imaged using a confocal microscope.

9.1.8 Primer Sequences

Table 9.6: Forward and Reverse Primer Sequences of Indicated Genes

Gene	Forward Primer (5' → 3')	Reverse Primer (5' → 3')	Product Size
36b4	AGATGCAGCAGATCCGCAT	GTTCTTGCCCATCAGCACC	59
Alb	GCTCGTCTGAGCCAGACATT	CAAGTTCCGCCCTGTCATCT	133
Ck19	GGGGGTTCAGTACGCATTGG	GAGGACGAGGTCACGAAGC	113
Cre	GCATTTCTGGGGATTGCTTA	ATTCTCCCACCGTCAGTACG	95
F4/80	TGGGATGCATAATCGCTGCT	CCTCAGAACCCACAGTGTCC	124
Gapdh	AGGTCGGTGTGAACGGATTTG	TGTAGACCATGTAGTTGAGGTCA	123
Gfap	CGGAGACGCATCACCTCTG	TGGAGGAGTCATTCGAGACAA	120
Tdo	ATGAGTGGGTGCCCGTTTG	GGCTCTGTTTACACCAGTTTGAG	105
Vim	CGTCCACACGCACCTACAG	GGGGGATGAGGAATAGAGGCT	74
vWF	CTTCTGTACGCCTCAGCTATG	GCCGTTGTAATTCCCACACAAG	125
Xbp1	ACACGCTTGGGAATGGACAC	CCATGGGAAGATGTTCTGGG	145, 171

9.1.9 Antibodies and Dilutions

Table 9.7: Antibodies with Corresponding Dilution and Supplier Information

Antibody	Host Species	Type	Applications	Supplier Information, Catalog Number
anti-Mouse	Goat	Secondary, HRP	WB - 1:5000	Bio-Rad, 1721011
anti-Mouse, Dylight 488	Goat	Secondary, Fluorescence	IF - 1:500	Thermo Fisher Scientific, 35503
anti-Mouse, Dylight 594	Goat	Secondary, Fluorescence	IF - 1:500	Thermo Fisher Scientific, 35511
anti-Rabbit	Goat	Secondary, HRP	WB - 1:5000	Thermo Fisher Scientific, 31460
anti-Rabbit, Dylight 488	Goat	Secondary, Fluorescence	IF - 1:500	Thermo Fisher Scientific, 35553
anti-Rabbit, Dylight 594	Goat	Secondary, Fluorescence	IF - 1:500	Thermo Fisher Scientific, 35561
BAX	Rabbit	Primary	WB - 1:5000	Abcam, ab32503
Beta-actin	Rabbit	Primary	WB - 1:5000	Cell Signaling Technology, 8457
DNA-RNA Hybrid, Clone S9.6	Mouse	Primary	DB - 1:5000	Abcam, ab256361
ds DNA	Mouse	Primary	DB - 1:5000	Abcam, ab27156
eIF2α	Rabbit	Primary	WB - 1:5000	Cell Signaling Technology, 9722
Hnf4a	Mouse	Primary	IF - 1:500	Abcam, ab41898
Hnf4a	Rabbit	Primary	IF - 1:500	Cell Signaling Technology, 3113
Ki67	Mouse	Primary	IHC - 1:250	BD Biosciences, 550609
p53	Rabbit	Primary	WB - 1:5000	Abcam, ab131442
phospho-eIF2α (S51)	Rabbit	Primary	WB - 1:5000	Cell Signaling Technology, 9721

*** WB – Western Blotting, DB – Dot Blotting, IF – Immunofluorescence, IHC – Immunohistochemistry

Table 9.7 (Continued): Antibodies with Corresponding Dilution and Supplier Information

Antibody	Host Species	Type	Applications	Supplier Information, Catalog Number
phospho-gamma-H2A.X (S139)	Rabbit	Primary	IF - 1:500	Abcam, ab11174
phospho-p53 (S15)	Rabbit	Primary	WB - 1:5000	Abcam, ab1431
RIP	Rabbit	Primary	WB - 1:5000	Abcam, ab106393
SRSF1	Rabbit	Primary	IF - 1:500, WB - 1:5000	Abcam, ab129108
SRSF1 (103)	Mouse	Primary	IF - 1:500, WB - 1:5000	Invitrogen, 32-4600
TBP	Mouse	Primary	WB - 1:10000	Thermo Fisher Scientific, MA5-14739

*** WB – Western Blotting, DB – Dot Blotting, IF – Immunofluorescence, IHC – Immunohistochemistry

9.2 Datasets

The following table provides details of the RNA-Seq data generated and analyzed for the current study. All raw RNA-seq data files are available for download from NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) under accession numbers GSE147005.

Table 9.8: SRSF1 Knockout Model RNA-seq Sample Information

GEO Accession Number	Model	Genotype	Treatment	Timepoint	Replicate	Total Number of Reads
GSM4412193	SRSF1 HKO	AlbCre +/-	NA	10 days	1	101964963
GSM4412194	SRSF1 HKO	AlbCre +/-	NA	10 days	2	110546221
GSM4412195	SRSF1 HKO	SRSF1 fl/fl; AlbCre +/-	NA	10 days	1	116181095
GSM4412196	SRSF1 HKO	SRSF1 fl/fl; AlbCre +/-	NA	10 days	2	117337934
GSM4412197	SRSF1 HKO	AlbCre +/-	NA	5 weeks	1	117680056
GSM4412198	SRSF1 HKO	AlbCre +/-	NA	5 weeks	2	116619589
GSM4412199	SRSF1 HKO	SRSF1 fl/fl; AlbCre +/-	NA	5 weeks	1	115650894
GSM4412200	SRSF1 HKO	SRSF1 fl/fl; AlbCre +/-	NA	5 weeks	2	117418842
GSM4412201	acSRSF1 HKO	SRSF1 fl/fl	AAV8-TBG-GFP	2 weeks	1	127573517
GSM4412202	acSRSF1 HKO	SRSF1 fl/fl	AAV8-TBG-GFP	2 weeks	2	176265351
GSM4412203	acSRSF1 HKO	SRSF1 fl/fl	AAV8-TBG-GFP	2 weeks	3	139425979
GSM4412204	acSRSF1 HKO	SRSF1 fl/fl	AAV8-TBG-iCRE	2 weeks	1	119934148
GSM4412205	acSRSF1 HKO	SRSF1 fl/fl	AAV8-TBG-iCRE	2 weeks	2	139028073
GSM4412206	acSRSF1 HKO	SRSF1 fl/fl	AAV8-TBG-iCRE	2 weeks	3	115531464

The following table provides details of the ENCODE datasets analyzed for the investigations in *Chapter 7*. These datasets are publicly available from the web portal at www.encodeproject.org.

Table 9.9: SR Protein Knockdown in HepG2 ENCODE RNA-seq Data Information

Experiment Target	Experiment Accession	Library-type	Utilized File Formats
SRSF7-human	ENCSR017PRS	paired-end	fastq, bam, bigWig
SRSF1-human	ENCSR094KBY	paired-end	fastq, bam, bigWig
SRSF9-human	ENCSR597XHH	paired-end	fastq, bam, bigWig
SRSF3-human	ENCSR376FGR	paired-end	fastq, bam, bigWig
SRSF5-human	ENCSR447UCG	paired-end	fastq, bam, bigWig
SRSF5-human	ENCSR781YNI	paired-end	fastq, bam, bigWig
Non-specific target control-human	ENCSR603TCV	paired-end	fastq, bam, bigWig
Non-specific target control-human	ENCSR264TUE	paired-end	fastq, bam, bigWig
Non-specific target control-human	ENCSR042QTH	paired-end	fastq, bam, bigWig

Chapter 10: References

Anczuków, Olga, Avi Z Rosenberg, Martin Akerman, Shipra Das, Lixing Zhan, Rotem Karni, Senthil K Muthuswamy, and Adrian R Krainer. 2012. “The Splicing Factor SRSF1 Regulates Apoptosis and Proliferation to Promote Mammary Epithelial Cell Transformation.” *Nature Structural & Molecular Biology* 19 (2): 220–28. https://doi.org/10.1038/nsmb.2207.

Aranda, Patrick S, Dollie M LaJoie, and Cheryl L Jorcyk. 2012. “Bleach Gel: A Simple Agarose Gel for Analyzing RNA Quality.” *Electrophoresis* 33 (2): 366–69. https://doi.org/10.1002/elps.201100335.

Asrih, Mohamed, and François R. Jornayvaz. 2015. “Metabolic Syndrome and Nonalcoholic Fatty Liver Disease: Is Insulin Resistance the Link?” *Molecular and Cellular Endocrinology*. Elsevier. https://doi.org/10.1016/j.mce.2015.02.018.

Belotserkovskii, Boris P, Jane Hae, Soo Shin, and Philip C Hanawalt. 2017. “Strong Transcription Blockage Mediated by R-Loop Formation within a G-Rich Homopurine-Homopyrimidine Sequence Localized in the Vicinity of the Promoter.” *Nucleic Acids Research* 45 (11): 6589–99. https://doi.org/10.1093/nar/gkx403.

Ben-Moshe, Shani, and Shalev Itzkovitz. 2019. “Spatial Heterogeneity in the Mammalian Liver.” *Nature Reviews Gastroenterology and Hepatology*. https://doi.org/10.1038/s41575-019-0134-x.

Bolger, Anthony M, Marc Lohse, and Bjoern Usadel. 2014. “Genome Analysis Trimmomatic: A Flexible Trimmer for Illumina Sequence Data” 30 (15): 2114–20.

https://doi.org/10.1093/bioinformatics/btu170.

Bray, Nicolas L, Harold Pimentel, Páll Melsted, and Lior Pachter. 2016. “Near-Optimal Probabilistic RNA-Seq Quantification.” *Nature Biotechnology* 34 (5): 525–27. https://doi.org/10.1038/nbt.3519.

Chen, Qing Rong, Rosemary Braun, Ying Hu, Chunhua Yan, Elizabeth M. Brunt, Daoud Meerzaman, Arun J. Sanyal, and Kenneth Buetow. 2013. “Multi-SNP Analysis of GWAS Data Identifies Pathways Associated with Nonalcoholic Fatty Liver Disease.” *PLoS ONE* 8 (7): 1–11. https://doi.org/10.1371/journal.pone.0065982.

Cyphert, T. J., A. L. Suchanek, B. N. Griffith, and L. M. Salati. 2013. “Starvation Actively Inhibits Splicing of Glucose-6-Phosphate Dehydrogenase MRNA via a Bifunctional ESE/ESS Element Bound by HnRNP K.” *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* 1829 (9): 905–15. https://doi.org/10.1016/j.bbagrm.2013.04.009.

Dara, Lily, Cheng Ji, and Neil Kaplowitz. 2011. “The Contribution of Endoplasmic Reticulum Stress to Liver Diseases.” *Hepatology* 53 (5): 1752–63. https://doi.org/10.1002/hep.24279.

Das, Shipra, Olga Anczuków, Martin Akerman, and Adrian R. Krainer. 2012. “Oncogenic Splicing Factor SRSF1 Is a Critical Transcriptional Target of MYC.” *Cell Reports* 1 (2): 110–17. https://doi.org/10.1016/j.celrep.2011.12.001.

Das, Shipra, and Adrian R Krainer. 2014. “Emerging Functions of SRSF1, Splicing Factor and Oncoprotein, in RNA Metabolism and Cancer.”

https://doi.org/10.1158/1541-7786.MCR-14-0131.

Davidson, Nicholas O., and Gregory S. Shelness. 2000. “Apolipoprotein B: MRNA Editing, Lipoprotein Assembly, and Presecretory Degradation.” *Annual Review of Nutrition*. Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA. https://doi.org/10.1146/annurev.nutr.20.1.169.

Dobin, Alexander, Carrie A Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R Gingeras. 2013. “Sequence Analysis STAR: Ultrafast Universal RNA-Seq Aligner” 29 (1): 15–21. https://doi.org/10.1093/bioinformatics/bts635.

Dong, Gaochao, Qixing Mao, Wenjie Xia, Youtao Xu, Jie Wang, Lin Xu, and Feng Jiang. 2016. “PKM2 and Cancer: The Function of PKM2 beyond Glycolysis (Review).” *Oncology Letters*. Spandidos Publications. https://doi.org/10.3892/ol.2016.4168.

Duncan, Andrew W. 2013. “Aneuploidy, Polyploidy and Ploidy Reversal in the Liver.” *Seminars in Cell and Developmental Biology*. Academic Press. https://doi.org/10.1016/j.semcdb.2013.01.003.

Duncan, Andrew W, Amy E Hanlon Newell, Weimin Bi, Milton J Finegold, Susan B Olson, Arthur L Beaudet, and Markus Grompe. 2012. “Aneuploidy as a Mechanism for Stress-Induced Liver Adaptation.” *Journal of Clinical Investigation* 122 (9): 3307–15. https://doi.org/10.1172/JCI64026.

Fagan-Solis, Katerina D., Dennis A. Simpson, Rashmi J. Kumar, Luciano G. Martelotto,

Lisle E. Mose, Naim U. Rashid, Alice Y. Ho, et al. 2020. “A P53-Independent DNA Damage Response Suppresses Oncogenic Proliferation and Genome Instability.” *Cell Reports*. https://doi.org/10.1016/j.celrep.2020.01.020.

Folch, Jordi, M Lees, and G H Sloane. 1957. “A Simple Method for the Isolation and Purification of Total Lipides from Animal Tissues.” *The Journal of Biological Chemistry* 226 (1): 497–509. http://www.jbc.org/.

Frevert, Ute, Sabine Engelmann, Sergine Zougbédé, Jörg Stange, Bruce Ng, Kai Matuschewski, Leonard Liebes, and Herman Yee. 2005. “Intravital Observation of Plasmodium Berghei Sporozoite Infection of the Liver.” Edited by Thomas Egwang. *PLoS Biology* 3 (6): e192. https://doi.org/10.1371/journal.pbio.0030192.

Galgani, J, and E Ravussin. 2008. “Energy Metabolism, Fuel Selection and Body Weight Regulation.” *International Journal of Obesity*. https://doi.org/10.1038/ijo.2008.246.

Gebhardt, Rolf. 1992. “Metabolic Zonation of the Liver: Regulation and Implications for Liver Function.” *Pharmacology and Therapeutics*. Pergamon. https://doi.org/10.1016/0163-7258(92)90055-5.

Goodman, Craig A., Danielle M. Mabrey, John W. Frey, Man Hing Miu, Enrico K. Schmidt, Philippe Pierre, and Troy A. Hornberger. 2010. “Novel Insights into the Regulation of Skeletal Muscle Protein Synthesis as Revealed by a New Nonradioactive in Vivo Technique.” *The FASEB Journal* 25 (3): 1028–39. https://doi.org/10.1096/fj.10-168799.

Goodman, Craig A, and Troy A Hornberger. 2013. “Measuring Protein Synthesis with SUnSET: A Valid Alternative to Traditional Techniques?” *Exercise and Sport Sciences Reviews* 41 (2): 107–15. https://doi.org/10.1097/JES.0b013e3182798a95.

Hir, Hervé Le, Jérôme Saulière, and Zhen Wang. 2016. “The Exon Junction Complex as a Node of Post-Transcriptional Networks.” *Nature Reviews Molecular Cell Biology*. Nature Publishing Group. https://doi.org/10.1038/nrm.2015.7.

Hocine, Sami, Robert H Singer, and David Grünwald. 2010. “RNA Processing and Export.” *Cold Spring Harbor Perspectives in Biology* 2 (12): a000752. https://doi.org/10.1101/cshperspect.a000752.

Howard, Jonathan M, and Jeremy R Sanford. 2015. “THE RNAissance Family: SR Proteins as Multifaceted Regulators of Gene Expression.” *Wiley Interdiscip Rev RNA* 6 (1): 93–110. https://doi.org/10.1002/wrna.1260.

Huang, Yingqun, Therese A Yario, and Joan A Steitz. 2004. “A Molecular Link between SR Protein Dephosphorylation and MRNA Export.” *Proceedings of the National Academy of Sciences* 101 (26): 9666–70. https://doi.org/10.1073/pnas.0403533101.

Hussain, M. Mahmood, Paul Rava, Meghan Walsh, Muhammad Rana, and Jahangir Iqbal. 2012. “Multiple Functions of Microsomal Triglyceride Transfer Protein.” *Nutrition and Metabolism*. BioMed Central. https://doi.org/10.1186/1743-7075-9-14.

Krainer, A R, G C Conway, and D Kozak. 1990. “Purification and Characterization of Pre-MRNA Splicing Factor SF2 from HeLa Cells.” *Genes & Development* 4 (7):

1158–71. https://doi.org/10.1101/gad.4.7.1158.

Kuleshov, Maxim V, Matthew R Jones, Andrew D Rouillard, Nicolas F Fernandez, Qiaonan Duan, Zichen Wang, Simon Koplev, et al. 2016. “Enrichr: A Comprehensive Gene Set Enrichment Analysis Web Server 2016 Update.” *Nucleic Acids Research* 44 (W1): W90–97. https://doi.org/10.1093/nar/gkw377.

Lee, Ursula E., and Scott L. Friedman. 2011. “Mechanisms of Hepatic Fibrogenesis.” *Best Practice and Research: Clinical Gastroenterology* 25 (2): 195–206. https://doi.org/10.1016/j.bpg.2011.02.005.

Lee, Yeon, and Donald C. Rio. 2015. “Mechanisms and Regulation of Alternative Pre-MRNA Splicing.” *Annual Review of Biochemistry*. https://doi.org/10.1146/annurev-biochem-060614-034316.

Li, Wan Chun, Kate L. Ralphs, and David Tosh. 2010. “Isolation and Culture of Adult Mouse Hepatocytes.” *Methods in Molecular Biology (Clifton, N.J.)*. https://doi.org/10.1007/978-1-59745-019-5_13.

Li, Xialu, and James L. Manley. 2005. “Inactivation of the SR Protein Splicing Factor ASF/SF2 Results in Genomic Instability.” *Cell* 122 (3): 365–78. https://doi.org/10.1016/j.cell.2005.06.008.

Li, Xialu, Tianhui Niu, and James L. Manley. 2007. “The RNA Binding Protein RNPS1 Alleviates ASF/SF2 Depletion-Induced Genomic Instability.” *RNA* 13 (12): 2108–15. https://doi.org/10.1261/rna.734407.

Liu, Cheng, Qing Tao, Mingyu Sun, Jim Z Wu, Wengang Yang, Ping Jian, Jinghua Peng, Yiyang Hu, Chenghai Liu, and Ping Liu. 2010. “Kupffer Cells Are Associated with Apoptosis, Inflammation and Fibrotic Effects in Hepatic Fibrosis in Rats.” *Laboratory Investigation* 90 (12): 1805–16. https://doi.org/10.1038/labinvest.2010.123.

Livak, K J, and T D Schmittgen. 2001. “Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2(-Delta Delta C(T)) Method.” *Methods (San Diego, Calif.)* 25 (4): 402–8. https://doi.org/10.1006/meth.2001.1262.

Long, Jennifer C., and Javier F. Caceres. 2009. “The SR Protein Family of Splicing Factors: Master Regulators of Gene Expression.” *The Biochemical Journal* 417 (1): 15–27. https://doi.org/10.1042/BJ20081501.

Love, Michael I., Wolfgang Huber, and Simon Anders. 2014. “Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2.” *Genome Biology* 15 (12): 550. https://doi.org/10.1186/s13059-014-0550-8.

Mabin, Justin W., Lauren A. Woodward, Robert D. Patton, Zhongxia Yi, Mengxuan Jia, Vicki H. Wysocki, Ralf Bundschuh, and Guramrit Singh. 2018. “The Exon Junction Complex Undergoes a Compositional Switch That Alters MRNP Structure and Nonsense-Mediated MRNA Decay Activity.” *Cell Reports* 25 (9): 2431-2446.e7. https://doi.org/10.1016/J.CELREP.2018.11.046.

Mah, L. J., A El-Osta, and T C Karagiannis. 2010. “Γh2AX: A Sensitive Molecular Marker of DNA Damage and Repair.” *Leukemia*. https://doi.org/10.1038/leu.2010.6.

Medina, Marisa Wong, Feng Gao, Devesh Naidoo, Lawrence L. Rudel, Ryan E. Temel, Allison L. McDaniel, Stephanie M. Marshall, and Ronald M. Krauss. 2011. “Coordinately Regulated Alternative Splicing of Genes Involved in Cholesterol Biosynthesis and Uptake.” Edited by Ying Xu. *PLoS ONE* 6 (4): e19420. https://doi.org/10.1371/journal.pone.0019420.

Michlewski, Gracjan, Jeremy R. Sanford, and Javier F. Cáceres. 2008. “The Splicing Factor SF2/ASF Regulates Translation Initiation by Enhancing Phosphorylation of 4E-BP1.” *Molecular Cell* 30 (2): 179–89. https://doi.org/10.1016/j.molcel.2008.03.013.

Morales, Julio C, Patricia Richard, Praveen L Patidar, Edward A Motea, Tuyen T Dang, James L Manley, and David A Boothman. 2016. “XRN2 Links Transcription Termination to DNA Damage and Replication Stress.” *PLoS Genetics* 12 (7). https://doi.org/10.1371/journal.pgen.1006107.

Müller-McNicoll, Michaela, Valentina Botti, Antonio M. de Jesus Domingues, Holger Brandl, Oliver D. Schwich, Michaela C. Steiner, Tomaz Curk, Ina Poser, Kathi Zarnack, and Karla M. Neugebauer. 2016. “SR Proteins Are NXF1 Adaptors That Link Alternative RNA Processing to MRNA Export.” *Genes and Development*. https://doi.org/10.1101/gad.276477.115.

Nostrand, Eric L. Van, Gabriel A. Pratt, Alexander A. Shishkin, Chelsea Gelboin-Burkhart, Mark Y. Fang, Balaji Sundararaman, Steven M. Blue, et al. 2016. “Robust Transcriptome-Wide Discovery of RNA-Binding Protein Binding Sites with Enhanced CLIP (ECLIP).” *Nature Methods* 13 (6): 508–14.

https://doi.org/10.1038/nmeth.3810.

Oslowski, Christine M, and Fumihiko Urano. 2011. *Measuring ER Stress and the Unfolded Protein Response Using Mammalian Tissue Culture System. Methods in Enzymology*. 1st ed. Vol. 490. Elsevier Inc. https://doi.org/10.1016/B978-0-12-385114-7.00004-0.

Pan, Qun, Ofer Shai, Leo J. Lee, Brendan J. Frey, and Benjamin J. Blencowe. 2008. "Deep Surveying of Alternative Splicing Complexity in the Human Transcriptome by High-Throughput Sequencing." *Nature Genetics*. https://doi.org/10.1038/ng.259.

Patterson, Ruth E., and Dorothy D. Sears. 2017. "Metabolic Effects of Intermittent Fasting." *Annual Review of Nutrition*. Annual Reviews Inc. https://doi.org/10.1146/annurev-nutr-071816-064634.

Poisson, Johanne, Sara Lemoinne, Chantal Boulanger, François Durand, Richard Moreau, Dominique Valla, and Pierre Emmanuel Rautou. 2017. "Liver Sinusoidal Endothelial Cells: Physiology and Role in Liver Diseases." *Journal of Hepatology*. https://doi.org/10.1016/j.jhep.2016.07.009.

Postic, Catherine, Masakazu Shiota, Kevin D. Niswender, Thomas L. Jetton, Yeujin Chen, J. Michael Moates, Kathy D. Shelton, Jill Lindner, Alan D. Cherrington, and Mark A. Magnuson. 1999. "Dual Roles for Glucokinase in Glucose Homeostasis as Determined by Liver and Pancreatic β Cell-Specific Gene Knock-Outs Using Cre Recombinase." *Journal of Biological Chemistry* 274 (1): 305–15. https://doi.org/10.1074/jbc.274.1.305.

Raabe, Martin, Laura M. Flynn, Constance H. Zlot, Jinny S. Wong, Murielle M. Véniant, Robert L. Hamilton, and Stephen G. Young. 1998. “Knockout of the Abetalipoproteinemia Gene in Mice: Reduced Lipoprotein Secretion in Heterozygotes and Embryonic Lethality in Homozygotes.” *Proceedings of the National Academy of Sciences of the United States of America* 95 (15): 8686–91. https://doi.org/10.1073/pnas.95.15.8686.

Raudvere, Uku, Liis Kolberg, Ivan Kuzmin, Tambet Arak, Priit Adler, Hedi Peterson, and Jaak Vilo. 2019. “G:Profiler: A Web Server for Functional Enrichment Analysis and Conversions of Gene Lists (2019 Update).” *Nucleic Acids Research* 47 (W1): W191–98. https://doi.org/10.1093/nar/gkz369.

Riepe, Celeste, Elena Zelin, Stacia Wyman, David Nguyen, Jin Rui Liang, Phillip Frankino, Zuriah Meacham, et al. 2018. “Double Stranded DNA Breaks and Genome Editing Trigger Ribosome Remodeling and Translational Shutdown.” *BioRxiv*, December, 486704. https://doi.org/10.1101/486704.

Rock, Kenneth L., and Hajime Kono. 2008. “The Inflammatory Response to Cell Death.” *Annual Review of Pathology: Mechanisms of Disease* 3 (1): 99–126. https://doi.org/10.1146/annurev.pathmechdis.3.121806.151456.

Ruas, Jorge L., James P. White, Rajesh R. Rao, Sandra Kleiner, Kevin T. Brannan, Brooke C. Harrison, Nicholas P. Greene, et al. 2012. “A PGC-1α Isoform Induced by Resistance Training Regulates Skeletal Muscle Hypertrophy.” *Cell*. https://doi.org/10.1016/j.cell.2012.10.050.

Samali, Afshin, Una Fitzgerald, Shane Deegan, and Sanjeev Gupta. 2010. “Methods for Monitoring Endoplasmic Reticulum Stress and the Unfolded Protein Response.” *International Journal of Cell Biology* 2010: 830307. https://doi.org/10.1155/2010/830307.

Sanchez, R. I., and F. C. Kauffman. 2010. “Regulation of Xenobiotic Metabolism in the Liver.” In *Comprehensive Toxicology, Second Edition*, 9:109–28. Elsevier Inc. https://doi.org/10.1016/B978-0-08-046884-6.01005-8.

Sanford, Jeremy R., Nicola K. Gray, Karsten Beckmann, and Javier F. Cáceres. 2004. “A Novel Role for Shuttling SR Proteins in MRNA Translation.” *Genes and Development* 18 (7): 755–68. https://doi.org/10.1101/gad.286404.

Santos-Pereira, José M., and Andrés Aguilera. 2015. “R Loops: New Modulators of Genome Dynamics and Function.” *Nature Reviews Genetics*. Nature Publishing Group. https://doi.org/10.1038/nrg3961.

Seimetz, Joseph, Waqar Arif, Sushant Bangru, Mikel Hernaez, and Auinash Kalsotra. 2019. “Cell-Type Specific Polysome Profiling from Mammalian Tissues.” *Methods* 155 (February): 131–39. https://doi.org/10.1016/J.YMETH.2018.11.015.

Sekine, Shigeki, Reiko Ogawa, Rie Ito, Nobuyoshi Hiraoka, Michael T. McManus, Yae Kanai, and Matthias Hebrok. 2009. “Disruption of Dicer1 Induces Dysregulated Fetal Gene Expression and Promotes Hepatocarcinogenesis.” *Gastroenterology* 136 (7): 2304-2315.e4. https://doi.org/10.1053/j.gastro.2009.02.067.

Sen, Supriya, Hassan Jumaa, and Nicholas J G Webster. 2013. “Splicing Factor SRSF3

Is Crucial for Hepatocyte Differentiation and Metabolic Function." *Nature Communications* 4 (May 2012): 1336. https://doi.org/10.1038/ncomms2342.

Shen, Shihao, Juw Won Park, Zhi Xiang Lu, Lan Lin, Michael D. Henry, Ying Nian Wu, Qing Zhou, and Yi Xing. 2014. "RMATS: Robust and Flexible Detection of Differential Alternative Splicing from Replicate RNA-Seq Data." *Proceedings of the National Academy of Sciences of the United States of America* 111 (51): E5593–5601. https://doi.org/10.1073/pnas.1419161111.

Si-Tayeb, Karim, Frédéric P Lemaigre, and Stephen a Duncan. 2010. "Organogenesis and Development of the Liver." *Developmental Cell*. https://doi.org/10.1016/j.devcel.2010.01.011.

Sollier, Julie, and Karlene A. Cimprich. 2015. "Breaking Bad: R-Loops and Genome Integrity." *Trends in Cell Biology*. Elsevier Ltd. https://doi.org/10.1016/j.tcb.2015.05.003.

Soneson, Charlotte, Michael I. Love, and Mark D. Robinson. 2015. "Differential Analyses for RNA-Seq: Transcript-Level Estimates Improve Gene-Level Inferences." *F1000Research* 4 (December): 1521. https://doi.org/10.12688/f1000research.7563.1.

Sundararaman, Balaji, Lijun Zhan, Steven M. Blue, Rebecca Stanton, Keri Elkins, Sara Olson, Xintao Wei, et al. 2016. "Resources for the Comprehensive Discovery of Functional RNA Elements." *Molecular Cell* 61 (6): 903–13. https://doi.org/10.1016/j.molcel.2016.02.012.

Tous, Cristina, and Andrés Aguilera. 2007. “Impairment of Transcription Elongation by R-Loops in Vitro.” *Biochemical and Biophysical Research Communications* 360 (2): 428–32. https://doi.org/10.1016/j.bbrc.2007.06.098.

Vernon, G, a Baranova, and Z M Younossi. 2011. “Systematic Review: The Epidemiology and Natural History of Non-Alcoholic Fatty Liver Disease and Non-Alcoholic Steatohepatitis in Adults.” *Alimentary Pharmacology & Therapeutics* 34 (3): 274–85. https://doi.org/10.1111/j.1365-2036.2011.04724.x.

Walsh, Callee M., Amanda L. Suchanek, Travis J. Cyphert, Alison B. Kohan, Wioletta Szeszel-Fedorowicz, and Lisa M. Salati. 2013. “Serine Arginine Splicing Factor 3 Is Involved in Enhanced Splicing of Glucose-6-Phosphate Dehydrogenase RNA in Response to Nutrients and Hormones in Liver.” *Journal of Biological Chemistry* 288 (4): 2816–28. https://doi.org/10.1074/jbc.M112.410803.

Weisend, Carla M, Jean A Kundert, Elena S Suvorova, Justin R Prigge, and Edward E Schmidt. 2009. “Cre Activity in Fetal AlbCre Mouse Hepatocytes: Utility for Developmental Studies.” *Genesis* 47 (12): 789–92. https://doi.org/10.1002/dvg.20568.

Wiśniewski, Jacek R, Alexandre Zougman, Nagarjuna Nagaraj, and Matthias Mann. 2009. “Universal Sample Preparation Method for Proteome Analysis.” *Nature Methods* 6 (5): 359–62. https://doi.org/10.1038/nmeth.1322.

Xu, Xiangdong, Dongmei Yang, Jian-Hua Ding, Wang Wang, Pao-Hsien Chu, Nancy D Dalton, Huan-You Wang, et al. 2005. “ASF/SF2-Regulated CaMKIIdelta Alternative

Splicing Temporally Reprograms Excitation-Contraction Coupling in Cardiac Muscle.” *Cell* 120 (1): 59–72. https://doi.org/10.1016/j.cell.2004.11.036.

Yu, Chi Yi, Elizabeth Theusch, Kathleen Lo, Lara M. Mangravite, Devesh Naidoo, Mariya Kutilova, and Marisa W. Medina. 2014. “HNRNPA1 Regulates HMGCR Alternative Splicing and Modulates Cellular Cholesterol Metabolism.” *Human Molecular Genetics*. https://doi.org/10.1093/hmg/ddt422.

Zahler, a M, W S Lane, J a Stolk, and M B Roth. 1992. “SR Proteins: A Conserved Family of Pre-MRNA Splicing Factors.” *Genes & Development* 6 (5): 837–47. https://doi.org/10.1101/gad.6.5.837.

Zhou, Lingyan, Congcong Li, Ling Gao, and Aihong Wang. 2015. “High-Density Lipoprotein Synthesis and Metabolism (Review).” *Molecular Medicine Reports*. Spandidos Publications. https://doi.org/10.3892/mmr.2015.3930.